

Document made available under the Patent Cooperation Treaty (PCT)

International application number: PCT/EP05/001688

International filing date: 18 February 2005 (18.02.2005)

Document type: Certified copy of priority document

Document details: Country/Office: DE
Number: 10 2004 013 843.5
Filing date: 20 March 2004 (20.03.2004)

Date of receipt at the International Bureau: 05 August 2005 (05.08.2005)

Remark: Priority document submitted or transmitted to the International Bureau in compliance with Rule 17.1(a) or (b)



World Intellectual Property Organization (WIPO) - Geneva, Switzerland
Organisation Mondiale de la Propriété Intellectuelle (OMPI) - Genève, Suisse

BUNDESREPUBLIK DEUTSCHLAND



Prioritätsbescheinigung über die Einreichung einer Patentanmeldung

Aktenzeichen: 10 2004 013 843.5

Anmeldetag: 20. März 2004


Anmelder/Inhaber: Degussa AG, 40474 Düsseldorf/DE

Bezeichnung: Expression von Nitrilhydratasen im Zwei-Vektor-Expressionssystem

IPC: C 12 N 15/70

Die angehefteten Stücke sind eine richtige und genaue Wiedergabe der ursprünglichen Unterlagen dieser Patentanmeldung.

München, den 21. Februar 2005
Deutsches Patent- und Markenamt
Der Präsident
Im Auftrag


Wehner

Expression von Nitrilhydratassen im Zwei-Vektor- Expressionssystem

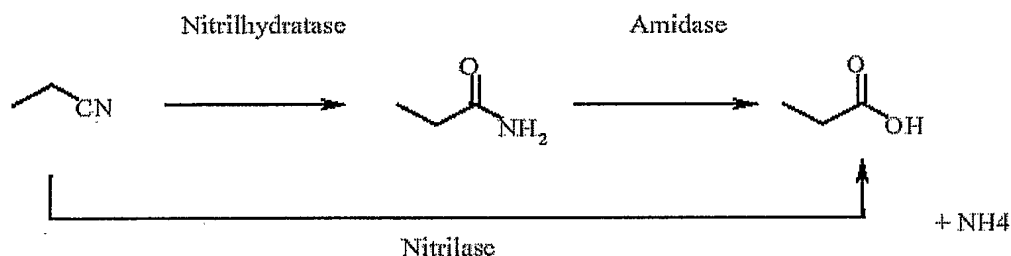
Die vorliegende Erfindung bezieht sich auf ein
Expressionssystem für die Herstellung von Nitrilhydratassen.

5 Nitrilhydratassen bestehen aus verschiedenen Untereinheiten.
Das gegenständliche System erlaubt die Herstellung der
Nitrilhydratassen in gegenüber dem Stand der Technik
verbesserter Art und Weise durch die getrennte Expression
der Nukleinsäuresequenzen kodierend für diese Untereinheiten
10 auf getrennten Plasmiden.

Die Strukturklassen der Amide und Carbonsäuren gewinnen mehr
und mehr an Bedeutung als Vorstufen von Feinchemikalien.
Spezielle Aminoamide und (proteinogene und nicht-
proteinogene) Aminosäuren sind Schlüsselintermediate für die
15 Synthese von pharmazeutischen und agrochemischen Produkten,
als auch im Lebensmittelbereich. Insbesondere
enantiomerenreine Amide und Aminosäuren spielen eine immer
größer werdende Rolle in den oben genannten
Anwendungsbereichen.

20 Aminonitril-Vorstufen, wie sie für die Herstellung der oben
angegebenen Verbindungsklassen benötigt werden, sind in
racemischer Form leicht über die so genannte
Streckersynthese zugänglich. Die so gewonnenen Nitrile
können anschließend mittels chemischer oder enzymatischer
25 Verseifung in die entsprechenden Amide und Carbonsäuren
überführt werden.

Es sind drei Enzyme bekannt, die an der enzymatischen
Hydrolyse von Nitrilen beteiligt sein können. Nitrilasen
setzen eine Nitril-Funktion direkt zur Säure um, wohingegen
30 Nitrilhydratassen (E.C. 4.2.1.84) hier das entsprechende Amid
bilden. Dieses kann durch eine Amidase (E.C. 3.5.1.4)
abschließend in die entsprechende Carbonsäure umgesetzt
werden (Schema 1).



Schema 1:

Die Verseifung von Nitrilen zu den entsprechenden Amiden und Säuren mittels isolierter Enzyme oder Ganz-Zell-Katalysatoren hilft große Mengen Salz zu sparen, welche ansonsten bei dem Neutralisierungsschritt nach der chemischen Verseifung von Nitrilen anfallen würden. Aus diesem Grund stellt die enzymatische Verseifung von Nitrilen zu z.B. Aminoamiden und/oder Aminosäuren ein nachhaltigeres Produktionsverfahren dar.

Nitrilhydratasen bestehen in ihrer aktiven Form aus 2 nicht-homologen α - und β -Untereinheiten. Diese bilden Heterodimere, Tetramere, und bei *Rhodococcus rhodochrous* J1 wurden sogar Decamere nachgewiesen. Die α - und β -Untereinheiten besitzen ungefähr die gleiche Größe, haben aber sonst keine Ähnlichkeiten untereinander. Nitrilhydratasen sind Metalloproteine die Fe^{3+} oder Co^{3+} enthalten (Bunch A. W. (1998), Nitriles, in: Biotechnology, Volume 8a, Biotransformations I, Chapter 6, Eds.: Rehm HJ, Reed G, Wiley-VCH, p. 277-324; Shearer J, Kung IY, Lovell S, Kaminsky W, Kovacs JA (2001) Why is there a "inert" metal center in the active site of nitrile hydratase? Reactivity and ligand dissociation from a five-coordinate Co(III) nitrile hydratase model. J Am Chem Soc 123: 463-468; Kobayashi M, Shimizu S (2000) Nitrile hydrolases. Current Opinion in Chemical Biology 4: 95-102).

Eine der größten Herausforderungen bisher ist die heterologe Darstellung von Nitrilhydratase in einem geeigneten Wirt, bevorzugt in *E. coli*. Dieses Gram negative Bakterium ist bekannt für seine hohen Expressionsraten heterologer Proteine. Ein weiterer Vorteil ist die Ausbeute an Biomasse in Hoch-Zelldichte-Fermentationen mit *E. coli*. Hierbei können Produktivitäten von über 100 g Biotrockenmasse (BTM) in 24 bis 44 Stunden erreicht werden (Lee SY (1996) High cell-density culture of *Escherichia coli*. TIBTECH 14:98-105; Riesenberg D, Guthke R (1999) High-cell-density cultivation of microorganisms. Appl Microbiol Biotechnol 51:422-430).

Die meisten Nitrilhydratase-Sequenzen der α - und β -Untereinheit sind aus der Gattung *Rhodococcus* bekannt. Aber gerade die Expression der Nitrilhydratase aus dieser Gattung in *E. coli* war bisher nur unter besonderen Schwierigkeiten möglich (Ikehata O, Nishiyama M, Horinouchi S, Beppu T (1989) Primary structure of nitrile hydratase deduced from the nucleotide sequence of a *Rhodococcus* species and its expression in *Escherichia coli*. Eur J Biochem 181: 563-570).

In der Literatur sind Ein-Vektor-Expressionssysteme für Nitrilhydratase beschrieben deren spezifische Aktivitäten zwischen 4,2 und 12,2 U/mg Gesamtprotein für Co-abhängige Nitrilhydratase aus *R. rhodochrous* J1 (Kobayashi M, Nishiyama M, Nagasawa T, Horinouchi S, Beppu T, Yamada H (1991) Cloning, nucleotide sequence and expression in *Escherichia coli* of two cobalt-containing nitrile hydratase genes from *Rhodococcus rhodochrous*. Biochim Biophys Acta 1129: 23-33) und 452 U/mg Gesamtprotein für eine eisen-abhängige Nitrilhydratase aus *Rhodococcus spec.* N-771 liegen (Njorri M, Yohda M, Odaka M, Matsushita Y, Tsujimura M, Yoshida T, Dohmae N, Takio K Endo I (1999) Functional expression of nitrile hydratase in *E. coli*: Requirement of a nitrile hydratase activator and a post-translational modification of a ligand cysteine. J Biochem 125: 696-704),

was ungefähr ca. 248 U/mg BTM (Biotrockenmasse) entspricht (Kalkulation nach Goodsell DS (1991) Inside a cell. TIBS 16: 203-206). Interessanterweise konnte die letztgenannte Aktivität mit Nitrilhydratasen aus *R. erythropolis*, welche
5 nahe verwandt sind mit *Rhodococcus spec.* N-711, mit ähnlichen Vektorsystemen und Anordnungen der Strukturgene nicht nachvollzogen werden. Es bestand daher immer noch ein Bedarf an Verfahren und Systemen welche es gestatten, die
10 ins Auge gefassten Enzyme in für technische Maßstäbe ausreichender Art und Weise zur Verfügung zu stellen.

Der Einsatz von Zwei-Vektor-Expressionssystemen für die heterologe Expression rekombinanter Proteine in *E. coli* ist dem Fachmann bereits bekannt, wie zum Beispiel die Bildung des motorischen Proteins Kinesin (Skowronek K, Kasprzak A
15 (2002) A two-plasmid system for independent genetic manipulation of subunits of homodimeric proteins and selective isolation of chimeric dimers. Analytical Biochemistry 300: 185-191), des Plaminogen-Proaktivator Streptokinase (Yazdani SS, Mukherjee KJ (2002) Continuous-
20 culture studies on the stability and expression of recombinant streptokinase in *Escherichia coli*; stability and expression of streptokinase in continuous culture. Bioprocess and Biosystems Engineering 24(6): 341-346), des Komplexes zweier humaner Proteine (hematopoietic cell
25 tyrosine phosphatase und der mitogen proteine kinase; Kholod N, Mustelin T (2001) Novel vectors for co-expression of two proteins in *E. coli*. 31: 322-328) oder der humanen Kreatin Kinase CKMB (WO95/12662) zeigen.

Bisher wurden jedoch noch keine heteromeren Enzyme die als
30 Biokatalysatoren in der chemischen Industrie eingesetzt werden, wie z.B. die Nitrilhydratasen, mit einem solchen System exprimiert.

Aufgabe war es daher ein Expressionssystem zu entwickeln, das es erlaubt, effizient sowohl cobalt- als auch
35 eisenabhängige Nitrilhydratasen aktiv in *E. coli* zu

exprimieren. Insbesondere sollte das erfindungsgemäße System in der Lage sein, die ins Auge gefassten Enzyme in gegenüber dem Stand der Technik erhöhter Expressionsrate und ggf. stabileren Formen zur Verfügung zu stellen, um so deren Einsatz im technischen Maßstab unter ökologischen und ökonomischen Gesichtspunkten vorteilhaft zu gestalten.

Diese und weitere nicht näher spezifizierte sich jedoch aus dem Stand der Technik in naheliegenderweise ergebende Aufgaben werden durch die Angabe eines Expressionssystems mit den Merkmalen des gegenständlichen Anspruchs 1 gelöst. Ansprüche 2 bis 8 beziehen sich auf bevorzugte Ausführungsformen des erfindungsgemäßen Expressionssystems. Ansprüche 9 und 10 sind auf Verfahren zur Herstellung von Nitrilhydratase bzw. (Amino-)Carbonsäuren oder (Amino-) Carbonsäureamide gerichtet. Anspruch 11 schützt einen mit dem Expressionssystem ausgestatteten Wirtsorganismus.

Dadurch, dass bei einem Expressionssystem für die gleichzeitige Expression der Nukleinsäuresequenzen kodierend für die verschiedenen Untereinheiten einer Nitrilhydratase das Expressionssystem mindestens je ein Plasmid mit mindestens einer Nukleinsäuresequenz kodierend für die jeweilige Untereinheit aufweist, gelangt man äußerst vorteilhaft und nichts desto weniger völlig überraschend zur Lösung der gestellten Aufgabe. Mit dem vorgeschlagenen Expressionssystem ist es möglich, die heterologe Expression der ins Auge gefassten Nukleinsäuresequenzen in für technische Maßstäbe ausreichender Art und Weise zu bewerkstelligen. Es kann dabei besonders überraschen, dass allein die getrennte Expression der eigentlich in einem Operon organisierten Nukleinsäuresequenzen kodierend für die entsprechenden Untereinheiten der Nitrilhydratase auf verschiedenen Plasmiden dazu beiträgt, die Aktivität der erhaltenen Nitrilhydratase um den Faktor > 8 gegenüber der „normalen“ Expression zu steigern. Dies war so aus dem Stand der Technik in naheliegender Weise nicht herleitbar.

Das erfindungsgemäße Expressionssystem kann in allen dem Fachmann für den vorliegenden Zweck in Frage kommenden Wirtsorganismen eingesetzt werden. Als Mikroorganismen sind diesbezüglich Organismen wie z.B. Hefen wie *Hansenula polymorpha*, *Pichia sp.*, *Saccharomyces cerevisiae*, Prokaryonten, wie *E. coli*, *Bacillus subtilis* oder Eukaryonten, wie Säugerzellen, Insektenzellen oder Pflanzenzellen zu nennen. Wirtsorganismen, in die die Nukleinsäuresequenzen aufweisende Plasmide kloniert werden können, dienen zur Vermehrung und Gewinnung einer ausreichenden Menge des rekombinanten Enzyms. Die Verfahren hierfür sind dem Fachmann wohlbekannt (Sambrook, J.; Fritsch, E. F. und Maniatis, T. (1989), *Molecular cloning: a laboratory manual*, 2nd ed., Cold Spring Harbor Laboratory Press, New York). Vorzugsweise sind *E. coli*-Stämme für diesen Zweck zu benutzen. Ganz besonders bevorzugt sind: *E. coli* XL1 Blue, NM 522, JM101, JM109, JM105, RR1, DH5 α , TOP 10-, HB101, BL21 codon plus, BL21 (DE3) codon plus, BL21, BL21 (DE3), MM294. Plasmide, mit denen das die erfindungsgemäße Nukleinsäure aufweisende Genkonstrukt vorzugsweise in den Wirtsorganismus kloniert wird, sind dem Fachmann ebenfalls bekannt (s.a. PCT/EP03/07148; s.u.). Ganz besonders bevorzugt ist ein Expressionssystem, das in *E. coli* BL21 als Wirt vorliegt.

Promotoren sind DNA-Sequenzbereiche, von denen aus die Transkription eines Gens oder Operons gesteuert wird. Die für die Ausführung der Erfindung besonders vorteilhaften Promotoren, welche insbesondere in *E. coli* einzusetzen sind, sind dem Fachmann bekannt (Sambrook, J.; Fritsch, E. F. und Maniatis, T. (1989), *Molecular cloning: a laboratory manual*, 2nd ed., Cold Spring Harbor Laboratory Press, New York). Es hat sich jetzt als vorteilhaft erwiesen, wenn die Expression der Nukleinsäuresequenzen kodierend für die Untereinheiten unter der Kontrolle von jeweils dem gleichen Promotor steht, damit die Nukleinsäuresequenzen kodierend für die Untereinheiten in möglichst gleicher Geschwindigkeit

exprimiert werden können. Geeignete Promotoren können solche ausgewählt aus der Gruppe T7, lac, tac, trp, ara oder rhamnose induzierbare sein. Weitere sind in (Cantrell, SA (2003) Vectors for the expression of recombinant proteins in E. coli. Methods in Molecular biology 235: 257-275; Sawers, G; Jarsch, M (1996) Alternative principles for the production of recombinant proteins in Escherichia coli. Applied Microbiology and Biotechnology 46(1): 1-9) genannt. Ganz besonders bevorzugt ist der Einsatz des so genannten T7-Promotors im erfindungsgemäßen Expressionssystem (Studier, W. F.; Rosenberg A. H.; Dunn J. J.; Dubendorff J. W.; (1990), Use of the T7 RNA polymerase to direct expression of cloned genes, Methods Enzymol. 185, 61-89; oder Broschüren der Firmen Novagen oder Promega).

Es hat sich als für die Funktionsfähigkeit des erfindungsgemäßen Expressionssystems nützlich erwiesen, dass auf den entsprechenden Plasmiden bestimmte Nukleinsäuresequenzen vorhanden sind, die für als Helferproteine bekannte Peptidsequenzen kodieren und deren Funktionen bisher weitgehend unbekannt sind. Diese sind dem Fachmann im Hinblick auf die Erzeugung aktiver Nitrilhydratasen bekannt (Nojiri M; Yohda M; Odaka M; Matsushita Y; Tsujimura M; Yoshida T; Dohmae N; Takio K; Endo I (1999) Functional expression of nitrile hydratase in Escherichia coli: requirement of a nitrile hydratase activator and post-translational modification of a ligand cysteine. Journal of biochemistry 125(4): 696-704). Ganz besonders bevorzugt ist ein Expressionssystem, bei dem pro eingesetztem Plasmidsatz mindestens eine Nukleinsäuresequenz kodierend für ein solches Helferprotein, insbesondere das p47K- (Seq. ID No. 33) oder p12K-Protein (Seq. ID No. 31), vorhanden ist. Ein Plasmidsatz bezeichnet dabei die Plasmide die erfindungsgemäß notwendig sind, eine aktive Nitrilhydratase aufzubauen.

Wie Eingangs näher erläutert sind Nitrilhydratasen aus verschiedenen Organismen bekannt (s.a. PCT/EP04/00338; Diss. s.o.). Vorzugsweise werden in dem erfindungsgemäßen Expressionssystem jedoch solche Nukleinsäuresequenzen verwendet, die für Untereinheiten von Nitrilhydratasen kodieren, welche ihren Ursprung in Nitrilhydratasen aus Rhodococcus-Stämmen haben. Die eingesetzten Nukleinsäuresequenzen können dabei gegenüber den Ursprungssequenzen aus Rhodococcus durch Mutagenese auf chemischer oder molekularbiologischer Basis verändert sein. Es kommen dabei insbesondere solche Nukleinsäuresequenzen in Betracht, die für Untereinheiten kodieren, die gegenüber den Wildtypsequenzen im Hinblick auf Aktivität und/oder Selektivität und/oder Stabilität verbessert sind. Die Verbesserung der Aktivität und/oder Selektivität und/oder Stabilität bedeutet erfindungsgemäß, dass die ins Auge gefassten Enzyme aktiver und/oder selektiver bzw. weniger selektiv oder unter den verwendeten Reaktionsbedingungen stabiler sind. Während die Aktivität und die Stabilität der Enzyme für die technische Anwendung naturgemäß möglichst hoch sein sollte, ist in Bezug auf die Selektivität dann von einer Verbesserung die Rede, wenn entweder die Substratselektivität abnimmt, die Enantioselektivität der Enzyme jedoch gesteigert ist.

Die Vorgehensweise zur Verbesserung der erfindungsgemäßen Nukleinsäuresequenzen bzw. der durch sie codierten Polypeptide durch Mutagenese-Methoden ist dem Fachmann hinlänglich bekannt. Als Mutagenese-Methoden kommen alle dem Fachmann für diesen Zweck zur Verfügung stehenden Methoden in Frage. Insbesondere sind dies die Sättigungsmutagenese, die Random-Mutagenese, in vitro-Rekombinations-Methoden sowie Site-Directed-Mutagenese (Eigen, M. und Gardiner, W. (1984), Evolutionary molecular engineering based on RNA replication, Pure Appl. Chem. 56, 967-978; Chen, K. und Arnold, F. (1991), Enzyme engineering for nonaqueous solvents: random mutagenesis to enhance activity of subtilisin E in polar organic media. Bio/Technology 9, 1073-

1077; Horwitz, M. und Loeb, L. (1986), Promoters Selected From Random DNA-Sequences, Proc Natl Acad Sci USA 83, 7405-7409; Dube, D. und L. Loeb (1989), Mutants Generated By The Insertion Of Random Oligonucleotides Into The Active-Site Of
5 The Beta-Lactamase Gene, Biochemistry 28, 5703-5707; Stemmer, P.C. (1994), Rapid evolution of a protein *in vitro* by DNA shuffling, Nature 370, 389-391 und Stemmer, P.C. (1994), DNA shuffling by random fragmentation and reassembly: *In vitro* recombination for molecular evolution.
10 Proc Natl Acad Sci USA 91, 10747-10751).
Besonders bevorzugt stammen die Nukleinsäuresequenzen kodierend für die Untereinheiten der Nitrilhydratasen aus Rhodococcus-Stämmen, insbesondere *R. erythropolis* 870-AN019.
In einer weiteren bevorzugten Ausführungsformen werden die
15 eingesetzten Nukleinsäuresequenzen dahingehend verändert, dass sie der „codon usage“ von *E. coli* besonders gut entsprechen. Es hat sich gezeigt, dass je mehr der Codon-Gebrauch des zu exprimierenden Gens dem von *E. coli* entspricht, die Ausbeuten der gewonnenen Enzyme weiter
20 gesteigert werden können. Besonders bevorzugt ist es deshalb, die Nukleinsäuresequenzen kodierend für die Untereinheiten der Nitrilhydratasen entsprechend der „codon usage“ von *E. coli* zu modifizieren. Unter „codon usage“ wird die Tatsache verstanden, dass unterschiedliche Organismen
25 unterschiedliche Basentriplets, welche für die gleichen Aminosäuren kodieren (Degeneration des genetischen Code), in unterschiedlicher Ausprägung gebrauchen.

Als Plasmide bzw. Vektoren kommen im Prinzip alle dem Fachmann für diesen Zweck zur Verfügung stehenden
30 Ausführungsformen in Frage. Derartige Plasmide und Vektoren können z. B. von Studier und Mitarbeiter (Studier, W. F.; Rosenberg A. H.; Dunn J. J.; Dubendorff J. W.; (1990), Use of the T7 RNA polymerase to direct expression of cloned genes, Methods Enzymol. 185, 61-89) oder den Broschüren der
35 Firmen Novagen, Promega, New England Biolabs, Clontech oder Gibco BRL entnommen werden. Weiter bevorzugte Plasmide und

- Vektoren können gefunden werden in: Glover, D. M. (1985), DNA cloning: a practical approach, Vol. I-III, IRL Press Ltd., Oxford; Rodriguez, R.L. und Denhardt, D. T (eds) (1988), Vectors: a survey of molecular cloning vectors and their uses, 179-204, Butterworth, Stoneham; Goeddel, D. V. (1990), Systems for heterologous gene expression, Methods Enzymol. 185, 3-7; Sambrook, J.; Fritsch, E. F. und Maniatis, T. (1989), Molecular cloning: a laboratory manual, 2nd ed., Cold Spring Harbor Laboratory Press, New York.
- 10 Plasmide, mit denen die die Nukleinsäuresequenzen aufweisenden Genkonstrukte kodierend für die Untereinheiten in ganz bevorzugter Weise in den Wirtsorganismus kloniert werden können, sind: pUC18/19 (Roche Biochemicals), pKK-177-3H (Roche Biochemicals), pBTac2 (Roche Biochemicals),
- 15 pKK223-3 (Amersham Pharmacia Biotech), pKK-233-3 (Stratagene) oder pET (Novagen). Ganz besonders bevorzugt ist ein Expressionssystem auf Basis von Plasmide der pET-Reihe. Äußerst bevorzugt ist der Einsatz von Plasmiden der gleichen Reihe sowohl für die Expression der
- 20 Nukleinsäuresequenz kodierend für die α - und β -Untereinheit.

- In einer weiteren Ausgestaltung bezieht sich die vorliegende Erfindung auf ein Verfahren zur Herstellung von Nitrilhydratase. Das Verfahren ist dadurch gekennzeichnet, dass es unter Verwendung eines wie oben dargestellten
- 25 erfindungsgemäßen Expressionssystems durchgeführt wird.

- In einer bevorzugten Ausführungsformen wird das erfindungsgemäße Verfahren so durchgeführt, dass die Expression bei Inkubationstemperaturen von kleiner 30 Grad Celsius, bevorzugt kleiner 25 Grad Celsius und ganz
- 30 besonders bevorzugt ≤ 20 Grad Celsius durchgeführt wird. Weiterhin vorteilhaft ist die Ausgestaltung, bei der während der Expression zum Medium Alkohole, insbesondere Ethanol, in einer Konzentration von kleiner 10% (w/w), bevorzugt kleiner 5% (w/w) und ganz besonders bevorzugt 2-4% (w/w) zugegeben
- 35 wird. Durch diese Maßnahmen wird erreicht, dass bei dem

erfindungsgemäßen Verfahren zur Herstellung der Nitrilhydratasen unlösliche Proteine (inclusion bodies), welche keine Aktivität entfalten, nicht oder nur in vermindertem Maße gebildet werden.

- 5 In einer nächsten Ausgestaltung bezieht sich die vorliegende Erfindung auf einen Wirtsorganismus aufweisend ein erfindungsgemäßes Expressionssystem. Wie weiter oben schon angedeutet können die Nukleinsäuresequenzen kodierend für die Untereinheiten von Nitrilhydratasen in wie oben
- 10 geschilderter erfindungsgemäßer Art und Weise in Plasmide integriert und in Wirtsorganismen transformiert werden. Zusätzlich kann neben dem erfindungsgemäßen Expressionssystem in dem Wirtsorganismus auch klonierte Gene für eine ggf. stereoselektiv arbeitende Amidase (z.B. die
- 15 aus WO2004/005517 oder EP 1318193) vorhanden sein. Ein so hergestellter Ganzzellkatalysator kann vorteilhafterweise beide am Nitril-Abbau beteiligten Enzyme produzieren, womit sichergestellt ist, dass das eingesetzte Nitril sofort zur entsprechenden Carbonsäuren umgewandelt wird.
- 20 Ganzzellkatalysatoren, welche mehrere an einer Reaktionskaskade beteiligte Enzyme enthalten, sind bereits bekannt (EP1216304). Deren Einsatz in der gegenständlichen Erfindung erfolgt in äquivalenter Art und Weise. Werden Rhamnose indizierbare Promotoren verwendet so sollte
- 25 ein Organismus wie in der DE10155928 genannt als Wirtsorganismus oder Ganzzellkatalysator eingesetzt werden. Weiterhin vorteilhaft ist der Einsatz eines *E. coli* BL21 codon plus, der ggf. gemäß dem der DE10155928 in äquivalenter Weise modifiziert ist.
- 30 Um die Expression der Enzyme im Hinblick auf ihre Umsetzungsgeschwindigkeiten abzustimmen, können die jeweils für die Nitrilhydratase und die Amidase codierenden Nukleinsäuresequenzen entsprechend ihren Umsetzungsraten in unterschiedliche Plasmide mit unterschiedlichen Kopienzahlen
- 35 kloniert und/oder unterschiedlich starke Promotoren für eine unterschiedlich starke Expression der Nukleinsäuresequenzen

verwendet werden. Bei derart abgestimmten Enzymsystemen tritt vorteilhafterweise eine Akkumulation einer ggf. inhibierend wirkenden Zwischenverbindung nicht auf und die betrachtete Reaktion kann in einer optimalen

5 Gesamtgeschwindigkeit ablaufen. Dies ist dem Fachmann jedoch hinlänglich bekannt (Gellissen, G.; Piontek, M.; Dahlems, U.; Jenzelewski, V.; Gavagan, J. W.; DiCosimo, R.; Anton, D. L.; Janowicz, Z. A. (1996), Recombinant *Hansenula polymorpha* as a biocatalyst. Coexpression of the spinach glycolate oxidase (GO) and the *S. cerevisiae* catalase T (CTT1) gene, Appl. Microbiol. Biotechnol. 46, 46-54; Farwick, M.; London, M.; Dohmen, J.; Dahlems, U.; Gellissen, G.; Strasser, A. W.; DE19920712).

15 Die Herstellung eines derartigen Ganzzellkatalysators ist dem Fachmann ebenfalls bekannt (Sambrook, J.; Fritsch, E. F. und Maniatis, T. (1989), Molecular cloning: a laboratory manual, 2nd ed., Cold Spring Harbor Laboratory Press, New York; Balbas, P. und Bolivar, F. (1990), Design and construction of expression plasmid vectors in *E. coli*, 20 Methods Enzymol. 185, 14-37; Rodriguez, R.L. und Denhardt, D. T (eds) (1988), Vectors: a survey of molecular cloning vectors and their uses, 205-225, Butterworth, Stoneham).

25 Eine weitere Ausgestaltung der gegenständlichen Erfindung bezieht sich auf ein Verfahren zur Herstellung von, ggf. enantiomerenangereicherten, (Amino-)Carbonsäuren oder (Amino-)Carbonsäureäureamiden. Auch dieses Verfahren wird unter Verwendung eines wie oben dargestellten Wirtsorganismus durchgeführt. Das bedeutet, dass die an der Herstellung der Carbonsäuren oder Carbonsäureamiden 30 beteiligten Nitrilhydratasen durch ein wie eben dargestellten Wirtsorganismus gewonnen werden.

Für die Anwendung können die betrachteten Enzyme in freier Form als homogen aufgereinigte Verbindungen oder als rekombinant (rec-) hergestelltes Enzym verwendet werden. 35 Weiterhin können die Enzyme auch als Bestandteil eines intakten Gastorganismus eingesetzt werden oder in Verbindung

mit der aufgeschlossenen und beliebig hoch aufgereinigten Zellmasse des Wirtsorganismus.

- Möglich ist ebenfalls die Verwendung der Enzyme in immobilisierter Form (Sharma B. P.; Bailey L. F. und Messing R. A. (1982), Immobilisierte Biomaterialien - Techniken und Anwendungen, Angew. Chem. 94, 836-852). Vorteilhafterweise erfolgt die Immobilisierung durch Lyophilisation (Paradkar, V. M.; Dordick, J. S. (1994), Aqueous-Like Activity of α -Chymotrypsin Dissolved in Nearly Anhydrous Organic Solvents, J. Am. Chem. Soc. 116, 5009-5010; Mori, T.; Okahata, Y. (1997), A variety of lipid-coated glycoside hydrolases as effective glycosyl transfer catalysts in homogeneous organic solvents, Tetrahedron Lett. 38, 1971-1974; Otamiri, M.; Adlercreutz, P.; Matthiasson, B. (1992), Complex formation between chymotrypsin and ethyl cellulose as a means to solubilize the enzyme in active form in toluene, Biocatalysis 6, 291-305). Ganz besonders bevorzugt ist die Lyophilisation in Gegenwart von oberflächenaktiven Substanzen, wie Aerosol OT oder Polyvinylpyrrolidon oder Polyethylenglycol (PEG) oder Brij 52 (Diethylenglycol-mono-cetylether) (Kamiya, N.; Okazaki, S.-Y.; Goto, M. (1997), Surfactant-horseradish peroxidase complex catalytically active in anhydrous benzene, Biotechnol. Tech. 11, 375-378).
- Äußerst bevorzugt ist die Immobilisierung an Eupergit® , insbesondere Eupergit C® und Eupergit 250L® (Röhm) (Eupergit.RTM. C, a carrier for immobilization of enzymes of industrial potential. Katchalski-Katzir, E.; Kraemer, D. M. Journal of Molecular Catalysis B: Enzymatic (2000), 10(1-3), 157-176.).
- Gleichfalls bevorzugt ist die Immobilisierung an Ni-NTA in Kombination mit dem His-Tag (Hexa-Histidin) ergänzten Polypeptid (Purification of proteins using polyhistidine affinity tags. Bornhorst, Joshua A.; Falke, Joseph J. Methods in Enzymology (2000), 326, 245-254). Die Verwendung als CLECs ist ebenfalls denkbar (St. Clair, N.; Wang, Y.-F.; Margolin, A. L. (2000), Cofactor-bound cross-linked enzyme crystals (CLEC) of alcohol dehydrogenase, Angew. Chem. Int.

Ed. 39, 380-383).

Durch diese Maßnahmen kann es gelingen aus Polypeptiden (Enzymen), welche durch organische Solventien instabil werden, solche zu generieren, die in Gemischen von wässrigen und organischen Lösungsmitteln bzw. ganz in Organik stabil sind und arbeiten können.

In den nachstehenden Experimenten werden folgende Stämme verwendet:

10 Liste der Verwendeten Stämme. (Brandão PFB, Clapp JP and Bull AT (2002). Discrimination and taxonomy of geographically diverse strains of nitrile-metabolising actinomycetes using chemometric and molecular sequencing techniques. Environmental Microbiology 4, 262-276; s.a. PCT/EP04/00338).

15 Tabelle 1:

Isolat	DSM	α -Untereinheit Seq. Nr.	β -Untereinheit Seq. Nr.
R. erythropolis 870-AN019	15258	1	3
R. erythropolis ENG-AN033	15261	5	7
R. erythropolis 871-AN042	15265	9	11
R. rhodochrous M8	Russian National Collection of Micro- organisms VKPM-S-926	13	15

Für die blunt-end Klonierung der Nitrilhydratasen und des p47K Proteins (Seq. ID No. 33) in pUC18/19 (xSmaI) (Abb. 1) aus den oben genannten Stämmen wurden folgende Primer benutzt:

Primer	Primersequenz	Amlifizierte Untereinheiten	Seq. Nr.
NH-Re-N	5'-GCC CGC ATA AGA AAA GGT GAA C	α , β , p47K	17
NH-Re-C-p47K	5'-GCA TGC CTT CAA ATC AGC CTG	α , β , p47K	18

5

Die ersten Expressionsexperimente im Ein-Vektor-Expressionssystem mit Nitrilhydratasen aus den *R. erythropolis*-Stämmen 870-AN019, 871-AN042 und ENG-AN033 wurden mit Plasmiden der pUC18/19 Reihe in verschiedenen *E. coli*-Stämmen durchgeführt. Um den optimalen Expressionswirt ermitteln zu können wurden die pUC18/19 Konstrukte mit den Nitrilhydratasen aus *R. erythropolis*-Stämmen 870-AN019, 871-AN042 und ENG-AN033 in verschiedene *E. coli*'s transformiert und unter der Kontrolle des *lac*-Promotors exprimiert. Es handelte sich dabei um folgende *E. coli*-Stämme: JM109, DH5 α , BL21 codon plus, BL21, HB101, MM294 und XL1blue.

Die Aktivitäten (Units pro g Zelltrockenmasse) der rekombinanten Nitrilhydratasen in den verschiedenen Wirten lassen sich wie folgt zusammenfassen: Es zeigte sich, dass die höchste Aktivität mit *E. coli* BL21 codon plus RIL (Firma Stratagen) erhalten wurde, in dem die Kopienzahl der für *E. coli* seltenen t-RNA's für Arginin (AGA/AGG), Isoleucin (AUA) und Leucin (CUA) erhöht vorliegen. Dieser Wirtsorganismus ist konstruiert worden, um speziell Gene mit einer an einen hohen GC-Gehalt angepassten „codon-usage“ zu exprimieren, wie z.B. die aus Rhodococcen (72% GC).

Die Nitrilhydratase aus *R. erythropolis* 870-AN019 in *E. coli* BL21 codon plus zeigt mit großem Abstand die höchste

Aktivität von 100 U/g BTM, gefolgt von Nitrilhydratase 870-AN019 in DH5 α (8 U/g BTM) und Nitrilhydratase ENG-AN033 in BL21 codon plus (7 U/g BTM).

- Die Aktivitäten aller anderen rekombinanten Organismen lagen unter 2 U/g BTM oder waren nicht nachweisbar. Unter optimierten Bedingungen wurde sogar für die Nitrilhydratase 870-AN019 eine Aktivität von 280 U/g BTM erreicht.

- Im folgenden wurde die Expression von eisenabhängigen Nitrilhydratase im Zwei-Vektor-Expressionssystem durchgeführt, wobei α - und β -Untereinheiten getrennt auf je einem Plamid vorlagen. Vorteil dieses Systems ist es, dass die beiden Untereinheiten jeweils direkt unter der Kontrolle des ggf. eingesetzten T7-Promoters liegen und somit die Transkripte für die Gene gleich stark gebildet werden. Das p47K-Helferprotein (Seq. ID No. 33) war von Fall zu Fall einem der beiden Untereinheiten nachgeschaltet. Als Expressionsvektoren dienten Plasmide der pET-Reihe (pET22b und pET26b) der Firma Stratagene. es wurde die Expression der Nitrilhydratase aus *R. erythropolis* 870-AN019 (Seq. Nr. 1 und 3) angestrebt.

Für die Klonierung der beiden Untereinheiten und des p47K Proteins (Seq. ID No. 33) wurden folgende Primer benutzt:

Primer	Primersequenz	Amplifizierter Orf	Seq. Nr.
NH019- α -for-Nde	5'-AGG GTG AAC CAT ATG TCA GTA ACG	α	19
NH019- α -rev-Bam	5'-TGT CGG ATC CAT CAG ACG GTG G	α	20
NH019- β -for-Nde	5'-AGC ACC ATA TGG ATG GAG TAC AC	β	21
NH019- β -rev-	5'-GTT GGG AAT TCA GGC	β	22

Eco	CGC AGG		
NH019-p47K- for-Bam	5'-CGC GGA TCC AAG AAG GAG ATA TAC ATG	p47K	23
NH019-p47K- rev-Hind	5'-CCG CAA CGT TCA AAC GGT CTG G	p47K	24

Für die α - und β -Untereinheit wurden Primer von der Nitrilhydratasesequenz aus *R. erythropolis* 870-AN019 abgeleitet (Brandao PFB, Clapp JP, Bull AT (2003) Diversity of nitrile hydratase and amidase enzyme genes in *Rhodococcus erythropolis* recovered from geographically distinct habitats. Applied and Environmental Microbiology 69(10): 5754-5766; s.a. PCT/EP04/00338), welche mit Schnittstellen für die Restriktionsendonukleasen *NdeI* (N-Terminal) bzw. *BamHI* (α -Untereinheit) oder *EcoRI* (β -Untereinheit) C-Terminal versehen waren. Die Primer für p47K wurden ebenfalls von diesem Organismus abgeleitet und enthielten die Schnittstellen für *BamHI* (N-terminal) bzw. *HindIII* (C-terminal). Die β -Untereinheit wurde in pET26b, die α -Untereinheit und p47K zusammen in pET22b kloniert. Es resultierten folgende Plasmide, welche in *E. coli* BL21 Codon (+) RIL transformiert wurden (siehe auch Abb. 2/3).

Verglichen wurden die Aktivitäten der Nitrilhydratase aus *R. erythropolis* 870-AN019 in den *E. coli*-Stämmen BL21 und BL21 codon plus RIL (siehe oben). Mit dem oben beschriebenen Expressionssystem wurden Aktivitäten von ca. 1750 U/g BTM mit *E. coli* BL21 erreicht und 2750 U/g BTM mit mit *E. coli* BL21 codon (+) (bezogen auf Benzonitril als Substrat). Dies bedeutet eine Steigerung gegenüber den Ein-Vektor-Expressionssystemen um den Faktor 5,3 bis 8,3.

50% des rekombinanten Proteins in der Zelle lagen hier allerdings als sogenannte „inclusion bodies“ vor. Zur weiteren Verbesserung der Aktivitäten der Nitrilhydratasen

wurden Parameter wie IPTG-Konzentration, Temperatur während der heterologen Expression und Zusatz von Additiven ins Medium untersucht. Die Verminderung der IPTG-Konzentration hatte keinen Einfluss auf die Löslichkeit der rekombinanten Nitrilhydratasen. Daher wurden alle weiteren Experimente mit 1 mM IPTG durchgeführt. Den größten Effekt auf die Verminderung der „inclusion body“-Bildung hatte die Absenkung der Inkubationstemperatur.

Durch die Reduktion der Temperatur bis auf 20°C nach Induktion der heterologen Expression konnte ein großer Anteil des Enzyms in die lösliche Fraktion überführt werden. Ein zusätzlichen positiven Effekt hatte die Zugabe von 3% Ethanol zum Medium. Unter diesen Bedingungen (1 mM IPTG, 20°C Inkubationstemperatur, 3% Ethanolzusatz) konnte eine Aktivität von 6480 U/g BTM erreicht werden.

Eine weitere Erhöhung der Aktivität wurde durch die Nutzung eines synthetischen Gens der Nitrilhydratase 870-AN019 erreicht. Hergestellt wurden die entsprechenden Nukleinsäuresequenzen kodierend für die α - und β -Untereinheiten unter Berücksichtigung der „codon usage“ von *E. coli* und unter Beibehaltung der Aminosäureabfolge der beiden Gene. Vorteil dieser synthetischen Gene sollte sein, dass die DNA-Sequenz optimal für eine Expression in *E. coli* angepasst ist und somit auch die Nutzung des sogenannten „codon plus“ Stammes entfallen kann. Die Untereinheiten wurden dann wie schon zuvor beschrieben getrennt in pET-Vektoren kloniert und unter den oben optimierten Bedingungen in *E. coli* BL21 exprimiert. Mit diesem Stamm wurde eine Aktivität von ca. 10.000 U/g BTM erreicht. Somit ist er mehr als doppelt so aktiv wie der Wildtyp *R. erythropolis* 870-AN019, der eine Aktivität von ca. 4760 U/g BTM mit Benzonitril als Substrat besitzt. Eine Zusammenfassung über die erreichten Aktivitäten gibt die Tabelle 2.

Die synthetischen Nitrilhydratase macht in dem *E. coli*-Wirt ca. 50% des gesamten gebildeten Zellproteins aus, wobei ca. 20% als gelöstes rekombinantes Protein vorliegen.

Tabelle 2: Übersicht über die gemessenen Nitrilhydratase-Aktivitäten mit den unterschiedlichen Expressionssystemen.

Expressions-system (Wirt & Promotor)	Expressionsbedingungen Nitrilhydratase aus 870-AN019		
	Native Untereinheiten, Standardmedium, 26°C	Native Untereinheiten, Medium + 3% EtOH, 20°C	Synthetische Untereinheiten Medium + 3% EtOH, 20°C
E. coli BL21 (DE3) & T7 Promoter	1750 U/g BTM	-	10080 U/g BTM
E. coli BL21 (DE3) codon usage & T7 Promoter	2750 U/g BTM	6480 U/g BTM	-

Abschließend wurde die Expression von cobaltabhängigen Nitrilhydratase-
 5 Nitrilhydratase im Zwei-Vektor-Expressionssystem untersucht. Die cobaltabhängige Nitrilhydratase aus *Rhodococcus rhodochrous* M8 (Pogorelva TE, Ryabchenko LE, Sunzov NI, Yanenko AS (1996) Cobalt-dependent transcription of nitrile hydratase gene in *Rhodococcus rhodochrous* M8.
 10 FEMS Microbiology Letters 144: 191-195) wurde kloniert und mit dem oben beschriebenen erfindungsgemäßen Expressionssystem hergestellt.

Die Sequenz der Nitrilhydratase (Seq. ID No. 13 und 15) ist bei GenBank (DNA Datenbank von DDBJ, EMBL und NCBI) unter der Kennung X86737 hinterlegt und frei zugänglich. Zur Klonierung der einzelnen Untereinheiten und des p12K-Proteins aus *R. rhodochrous* M8 (Seq. ID No. 31) wurden folgende Primer benutzt:

Primer	Primersequenz	Amlifizierte Untereinheiten	Seq. Nr.
M8- α -for-Nde	5'-AGG AAT ACG CAT ATG AGC GAGC ACG TC	α	25
M8- α -rev-Bam	5'-GTG TGG ATC CAC TCA TAC GAT CAC TTC CTG	α	26
M8- β -for-Nde	5'-AGG AAT GAG CAT ATG GAT GGT ATC CAC GAC A	β	27
M8- β -rev-Bam	5'-ATC GGG ATC CTT TCA CGC AGA GAT CAG GTA CGG	β	28
M8-p12K-for-Bam	5'-CTC AGG ATC CAA GGA GTG ATC GTA TGA GTG AAG AC	p12K	29
M8-p12K-rev-Sac	5'-ACA GGA GCT CTC AGT CGA TGA TGG CC	p12K	30

Für die α - und β -Untereinheit wurden Primer von der Nitrilhydratasesequenz aus M8 abgeleitet, welche mit Schnittstellen für die Restriktionsendonukleasen NdeI (N-Terminal) bzw. BamHI (C-Terminal) versehen waren. Die Primer für p12K wurden von der Sequenz von *R. rhodochrous* J1 abgeleitet und enthielten die Schnittstellen für BamHI (N-terminal) bzw. SacI (C-terminal). Die β -Untereinheit (Seq. ID No. 15) wurde in pET26b, die α -Untereinheit (Seq. ID No. 13) und p12K (Seq. ID No. 31) in pET22b kloniert. Es

resultierten folgende Plasmide, welche in *E. coli* BL21 Codon(+) RIL transformiert wurden (Abb. 4/5).

Im Gegensatz zur Expression von eisenabhängigen Nitrilhydratase in *E. coli*, musste im Fall der
5 cobaltabhängigen Nitrilhydratase die Zellen über mehrere Generationen vorkultiviert werden, um diese an das toxische Cobalt (0,5 mM eingesetzt) zu gewöhnen.

10 In Tabelle 3 sind die Aktivitäten der rekombinanten cobaltabhängigen (*R. rhodochrous* M8) und eisenabhängigen (*R. erythropolis* 870-AN019) Nitrilhydratase im Vergleich dargestellt.

Tabelle 3: Gemessene Aktivitäten der cobalt- und eisenabhängigen Nitrilhydratase nach erfindungsgemäßer Expression mit Acrylamid als Substrat.

Rekombinante Nitrilhydratase aus	Aktivität (U/mg); Substrat Acrylamid
<i>R. rhodochrous</i> M8 (Co)	160
<i>R. erythropolis</i> 870-AN019 (Fe) (synthetisch s.o.)	250

15

Damit konnte gezeigt werden, dass das erfindungsgemäße Expressionssystem sowohl für cobalt- als auch für eisenabhängige Nitrilhydratase im vorteilhafter Art und Weise zu gebrauchen ist. Es können insbesondere darmatische
20 Aktivitätssteigerungen erzielt werden, wenn die Expressionssysteme und/oder Sequenzen entsprechend optimiert werden. Durch die erfolgreiche getrennte Expression der Untereinheiten der Nitrilhydratase ist es darüber hinaus möglich neue Kombinationen von Untereinheiten verschiedener
25 Nitrilhydratase zu untersuchen, was die Diversifikation der

Nitrilhydratasen und damit deren Eigenschaften steigern hilft. Dies war so aus dem Stand der Technik zum Zeitpunkt der Erfindung in naheliegender Weise nicht ableitbar.

5 Unter optisch angereicherten (enantiomerenangereicherten, enantiomer angereicherten) Verbindungen wird im Rahmen der Erfindung das Vorliegen einer optischen Antipode im Gemisch mit der anderen in >50 mol-% verstanden.

10 Unter dem Begriff Nukleinsäuresequenzen werden alle Arten von einzelsträngiger oder doppelsträngiger DNA als auch RNA oder Gemische derselben subsumiert.

Die Organismen 870-AN019, ENG-AN033 und 871-AN042 wurden bei der Deutsche Sammlung für Mikroorganismen und Zellkulturen, Mascheroder Weg 4, 38124 Braunschweig gemäß dem Budapester Vertrag durch die Anmelderin am 22.10.2002 hinterlegt.

15 Der organismus *Rhodococcus rodochrous* M8 ist in der All-Russian National Collection of Microorganisms unter der Nummer VKPM-S-926 hinterlegt. Die entsprechenden Sequenzen können der Gendatenbank (s.o.) entnommen werden.

20 Der terminus Expressionssystem wird erfindungsgemäß so verstanden, dass es sich dabei um biologisches Material auf Nukleinsäurebasis handelt, welches im Stande ist, in Organismen die Expression der ihm innewohnenden Nukleinsäuresequenzen kodierend für die Nitrilhydratase-Unterheiten zu bewerkstelligen. Insbesondere sind dies
25 Plasmide und Vektoren.

Im Rahmen der Erfindung werden die Ausdrücke Protein und Polypeptid synonym benutzt.

Beschreibung der Abbildungen:

Abbildung 1: Die Vektorkarte zeigt die allgemeine Anordnung der α - und β -Untereinheiten der Nitrilhydratase bzw. des Orf's p47K im Plasmid pUC18 in einem Expressionssystem mit einem Vektor.

Abbildung 2/3: Die Vektorkarten zeigen die Anordnung der α - und β -Untereinheiten der Nitrilhydratase bzw. des Orf's p47K aus *R. erythropolis* 870-AN019 in Plasmiden der pET-Reihe in einem Expressionssystem mit zwei Vektoren.

10 Abbildung 4/5: Die Vektorkarten zeigen die Anordnung der α - und β -Untereinheiten bzw. des Orf's p12K aus *R. rhodochrous* M8 in Plasmiden der pET-Reihe in einem Expressionssystem mit zwei Vektoren.

Experimenteller Teil:

Kultivierung von Mikroorganismen

Die Kultivierung der *E. coli* Zellen und deren Aufbewahrung
5 erfolgte nach Sambrook et al. (Sambrook, J.; Fritsch, E. F.
und Maniatis, T. (1989), *Molecular cloning: a laboratory
manual*, 2nd ed., Cold Spring Harbor Laboratory Press, New
York).

Zur Kultivierung der Rhodococcen wurde ein
10 Chelatmineralmedium (CMM) nach Heald et al. 2001, (Heald S.
C., P. F. B. Brandao, R. Hardicre and A. T. Bull, 2001:
*Physiology, biochemistry and taxonomy of deep-sea nitrile
metabolising Rhodococcus strains*. Antonie van Leeuwenhoek
80, 169-183) eingesetzt. Das CM-Medium wird mit 5 g/l
15 Glucose supplementiert. Für Stämme mit cobaltabhängigen
Enzymen wird 2,4 mg/l $\text{CoCl}_2 \cdot 6 \text{H}_2\text{O}$, für Stämme mit
eisenabhängigen Enzymen 50 mg/l $\text{FeSO}_4 \cdot 7 \text{H}_2\text{O}$ eingesetzt.

Die aus der Kultur entnommene Probe wurden mit entsprechend
der Kultivierung eingesetztem Kaliumphosphatpuffer (6 ml/l
20 einer 200 g/l K_2HPO_4 -Lösung und 4 ml/l einer 157,5 g/l
 KH_2PO_4 -Lösung) so verdünnt, dass der Meßbereich zwischen
0,05 und 0,3 lag. Als Referenz diente der Puffer. Gemessen
wurde bei 600 nm.

PCR

25 Zur Amplifikation von DNA-Fragmenten aus Rhodococcen wurden
pro 50 µl Ansatz folgende Komponenten zusammenpipettiert:

100 ng Chromosomale DNA
1 µl dNTP -Mix (alle 10 mM)
5 µl Puffer
30 2 µl DMSO
0,5 µl Herculase (2,5 U)
ad 50 µl H_2O

Der Thermocycler wurde folgendermaßen programmiert:

- A. 3 min 98 °C
- B. 40 sec 98°C
- C. 40 sec X°C
- 5 D. Y min 72°C
- E. 5 min 72°C
- F. 4°C

- 10 Die Schritte B, C, D wurden 30 mal wiederholt. Die Annealingtemperatur X (C) errechnet sich aus der Schmelztemperatur der verwendeten Primer und die Inkubationsdauer Y (D) aus der Länge des zu amplifizierenden Genes (1 kb DNA gleich 1 Minute-Regel).

Verdau mit Restriktionsenzymen

- 15 Die zu schneidende DNA wird mit 5 U Restriktionsenzym und den zugehörigen Puffer versehen und wenn nicht anders erforderlich bei 37°C inkubiert. Der Verdau chromosomaler DNA erfolgt mit 10 U Enzym. Die Inkubationsdauer beträgt 1,5 - 2,5 Stunden.

20

Behandlung mit alkalischer Phosphatase

- 25 Um zu verhindern, dass Vektoren, welche nur mit einer Restriktionsendonuklease geschnitten wurden, mit sich selbst religieren, wird mit Hilfe der alkalischen Phosphatase der am 5'-Ende überhängende Phosphatrest entfernt. Nur durch Insertion eines DNA-Fragments kann wieder zirkuläre DNA entstehen.

- 30 Der mit einer Restriktionsendonuklease geschnittene Vektor wird 15 min bei 65°C inkubiert, um die Restriktionsendonuklease abzustoppen. Anschließend wird der Dephosphorylierungspuffer zugegeben und mit 1 U alkalischer

Phosphatase aus Schrimps wird 10 min bei 37°C inkubiert. Das Enzym wird durch eine anschließende Gelelektrophorese von der Vektor-DNA abgetrennt.

5 Behandlung mit T4-DNA-Ligase

Für die Ligation werden Vektor und Insert im Verhältnis 1:3 eingesetzt. Das Volumen wird möglichst gering gewählt (7-20 µl) Der Ansatz wird in Ligationspuffer und in Anwesenheit von 1 U Ligase bei 16 °C über Nacht inkubiert.

10 Transformation

Zum Ligationsansatz werden 100 µl kompetente Zellen pipettiert und der Ansatz durch wiederholtes Aufziehen der Pipette gemischt. Nach 30 min Inkubation auf Eis wird ein Hitzeschockschritt bei 42 °C für 45 sec durchgeführt und
15 wieder 2 min auf Eis inkubiert. Es werden 120-900 µl SOC-Medium zugegeben und der Ansatz wird 45 min bei 37°C unter Agitation inkubiert. Anschließend wird der Ansatz ausplattiert und bei 37°C über Nacht inkubiert.

20 Expression der verschiedenen Nitrilhydratasen (Tabelle 1) im Ein-Vektor-Expressionssystem

Folgendes Protokoll wurde zur Expression verwendet: 50 ml LB_{amp100}-Medium mit 2 mM Fe-Citrat wurde 1%ig mit einer Übernachtskultur angeimpft. Nach Erreichen einer OD₆₀₀ von ca.
25 0,5 wurde mit 1 mM IPTG (Isopropylthiogalactosid) die Expression der Nitrilhydratasen in den verschiedenen *E. coli*s induziert. Die Ernte der Zellen erfolgte ca. 24 Stunden nach Induktion. Die Expression der Nitrilhydratasen in Stamm DH5α erfolgte konstitutiv, da dieser Stamm den
30 *lac*-Repressor nicht überexprimiert. Dadurch entfällt hier der Induktionsschritt mit IPTG.

Aktivitätsnachweis mit Benzonitril als Substrat:

Die Biotransformation wurde im 10 ml Maßstab durchgeführt mit ca. 100 mg Biofeuchtmasse ($OD_{600} = 5$) im Kaliumphosphatpuffer (100 mM) pH7,0. Die Inkubation erfolgte bei 30°C und die Substratkonzentration betrug ca. 5 mM Benzonitril. Die Probennahme erfolgte alle 5 - 10 min über einen Zeitraum von maximal 1 Stunde. Das Probenvolumen betrug 100 µl und die Reaktion wurde durch Zugabe von 10 µl 50%ige Phosphorsäure gestoppt.

10 Die Konzentrationen von Benzonitril und Benzamid wurden dann mittels HPLC bestimmt:

Säule: RP18-Säule Phenomenex Hypersil ODS 5µ (mit Vorsäule)

Fließmittel: 10 mM K_2HPO_4 , (pH 2.3)

Flußrate: 1 ml / min

15 Wellenlänge: 202 nm

Injektionsvolumen: 20 µl

Dauer HPLC Lauf: 12-15 min

Die Berechnung der Aktivität erfolgte über die Kalkulation von einem µmol Umsatz nach einer Minute, was einem U (Unit entspricht). Spezifische Aktivitäten werden in U pro g BTM oder mg Protein angegeben.

Expression der Nitrilhydratase aus *R. erythropolis* 870-AN019 im Zwei-Vektor-Expressionssystem

Die Expression der Konstrukte mit T7-Promotoren erfolgte nach folgendem Protokoll:

25 50 ml LB_{amp100} -Medium mit 2 mM Fe-Citrat und jeweils 50 µg/ml Kanamycin & Ampicillin wurde 1%ig mit einer Übernachtskultur angeimpft. Nach Erreichen einer OD_{600} von ca. 0,5 wurde mit 1 mM IPTG (Isopropylthiogalactosid) die Expression der Nitrilhydratasen induziert. Die Ernte der Zellen erfolgte ca. 24 Stunden nach Induktion bei 26°C.

Expression der Nitrilhydratase aus *R. rhodochrous* M8 im Zwei-Vektor-Expressionssystem

Die Expression der Konstrukte mit T7-Promotoren erfolgte nach folgendem Protokoll:

- 50 ml LB_{amp100}-Medium mit 0,5 mM CoCl₂ und jeweils 50 µg/ml Kanamycin & Ampicillin wurde 1%ig mit einer Übernachtskultur angeimpft. Nach Erreichen einer OD₆₀₀ von ca. 0,5 wurde mit 1 mM IPTG (Isopropylthiogalactosid) die Expression der NHasen induziert. Das Medium wurde zusätzlich mit 3% (w/v) Ethanol versetzt. Die Ernte der Zellen erfolgte ca. 24 Stunden nach Induktion bei 26°C.

10 Aktivitätsbestimmung

- Die Biotransformation zur Aktivitätsbestimmung mit den rekombinanten Nitrilhydratasen *R. rhodochrous* M8 und *R. erythropolis* 870-AN019 in *E. coli* erfolgte in kleinem Maßstab. Die Biotransformation wird in einem 1,5 ml Eppendorfcup bei 20°C durchgeführt. Im Biotransformationsansatz wurde eine OD von 0,4 eingesetzt. 500 µl einer 4 % Acrylnitril-Lösung in 10 mM Kaliumphosphatpuffer, pH 7,5 sowie der Puffer werden bei 20 °C vorinkubiert. Die Reaktion wird durch Zugabe der Zellen gestartet. Dabei beträgt die Summe des Puffer- und Zellvolumens 500 µl. Sofort nach Mischung des Ansatzes werden 100 µl entnommen und zu 1,5 µl vorgelegter konzentrierter HCl pipettiert. Nach Mischung wird die Probe 2 min bei 13000 rpm in der Eppendorf-Zentrifuge abzentrifugiert und 70 µl des Überstands zur Analyse mittels HPLC aufbewahrt bei - 20°C. Die Probennahme erfolgte alle 5 - 10 min über einen Zeitraum von maximal 2 Stunden.

Analyse von Acrylnitril, Acrylamid und Acrylsäure

Mit der nachfolgend beschriebenen HPLC Methode ist es möglich Acrylnitril, Acrylamid und Acrylsäure in kurzer Zeit zu analysieren und die Konzentration dieser Substanzen zu bestimmen:

Säule: Synergi 4µ Hydro-RP mit Vorsäule

Fließmittel: 0,1 % H_3PO_4 in 10 % Acetonitril, 90 % H_2O

Flußrate: 0,5 ml / min

Wellenlänge: 202 nm

10 Injektionsvolumen: 5 µl

Dauer HPLC Lauf: 10 min, letzter Peak nach 5,5 min

Die Berechnung der Aktivität erfolgte über die Kalkulation von einem µmol Umsatz nach einer Minute, was einem U (Unit entspricht). Spezifische Aktivitäten werden in U pro g BTM oder mg Protein angegeben.

15

Patentansprüche:

1. Expressionssystem für die gleichzeitige Expression der Nukleinsäuresequenzen kodierend für die verschiedenen Untereinheiten einer Nitrilhydratase,
5 dadurch gekennzeichnet, dass
das Expressionssystem mindestens je ein Plasmid mit mindestens einer Nukleinsäuresequenz kodierend für die jeweilige Untereinheit aufweist.
2. Expressionssystem nach Anspruch 1,
10 dadurch gekennzeichnet, dass
dieses in *E. coli* als Wirt vorliegt.
3. Expressionssystem nach einem oder mehreren der vorhergehenden Ansprüche,
15 dadurch gekennzeichnet, dass
die Expression der Nukleinsäuresequenzen kodierend für die Untereinheiten unter der Kontrolle von jeweils dem gleichen Promotor steht.
4. Expressionssystem nach Anspruch 3,
20 dadurch gekennzeichnet, dass
der Promotor ein T7-Promotor ist.
5. Expressionssystem nach einem oder mehreren der vorhergehenden Ansprüche,
25 dadurch gekennzeichnet, dass
pro eingesetztem Plasmidsatz mindestens eine Nukleinsäuresequenz kodierend für das p47K- oder p12K-Protein vorhanden ist.
6. Expressionssystem nach einem oder mehreren der vorhergehenden Ansprüche,
30 dadurch gekennzeichnet, dass
die Nukleinsäuresequenzen kodierend für die Untereinheiten der Nitrilhydratasen aus *Rhodococcus*-Stämmen stammen.

7. Expressionssystem nach einem oder mehreren der vorhergehenden Ansprüche, dadurch gekennzeichnet, dass man die Nukleinsäuresequenzen kodierend für die Untereinheiten der Nitrilhydratasen entsprechend der „codon usage“ von *E. coli* modifiziert einsetzt.
5
8. Expressionssystem nach einem oder mehreren der vorhergehenden Ansprüche, dadurch gekennzeichnet, dass man als Plasmide solche der pET-Reihe benutzt.
10
9. Verfahren zur Herstellung von Nitrilhydratasen unter Verwendung eines Expressionssystems gemäß einem oder mehreren der Ansprüche 1 bis 8.
10. Wirtsorganismus aufweisend ein Expressionssystem gemäß einem oder mehreren der Ansprüche 1 bis 8.
15
11. Verfahren zur Herstellung von, ggf. enantiomerenangereicherten, (Amino-)Carbonsäuren oder (Amino-)Carbonsäureäureamiden unter Verwendung eines Wirtsorganismus gemäß einem oder mehreren der Ansprüche 1 bis 8.
20

Zusammenfassung:

Die vorliegende Erfindung bezieht sich auf ein neues
Expressionssystem für Nitrilhydratase. Die aus
Untereinheiten aufgebauten Nitrilhydratasen werden dabei in
5 der Art und Weise gebildet, dass die jeweilige
Untereinheiten auf unterschiedlichen Plasmiden beheimatet
sind und gleichzeitig in *E. coli* exprimiert werden.

Abb. 1

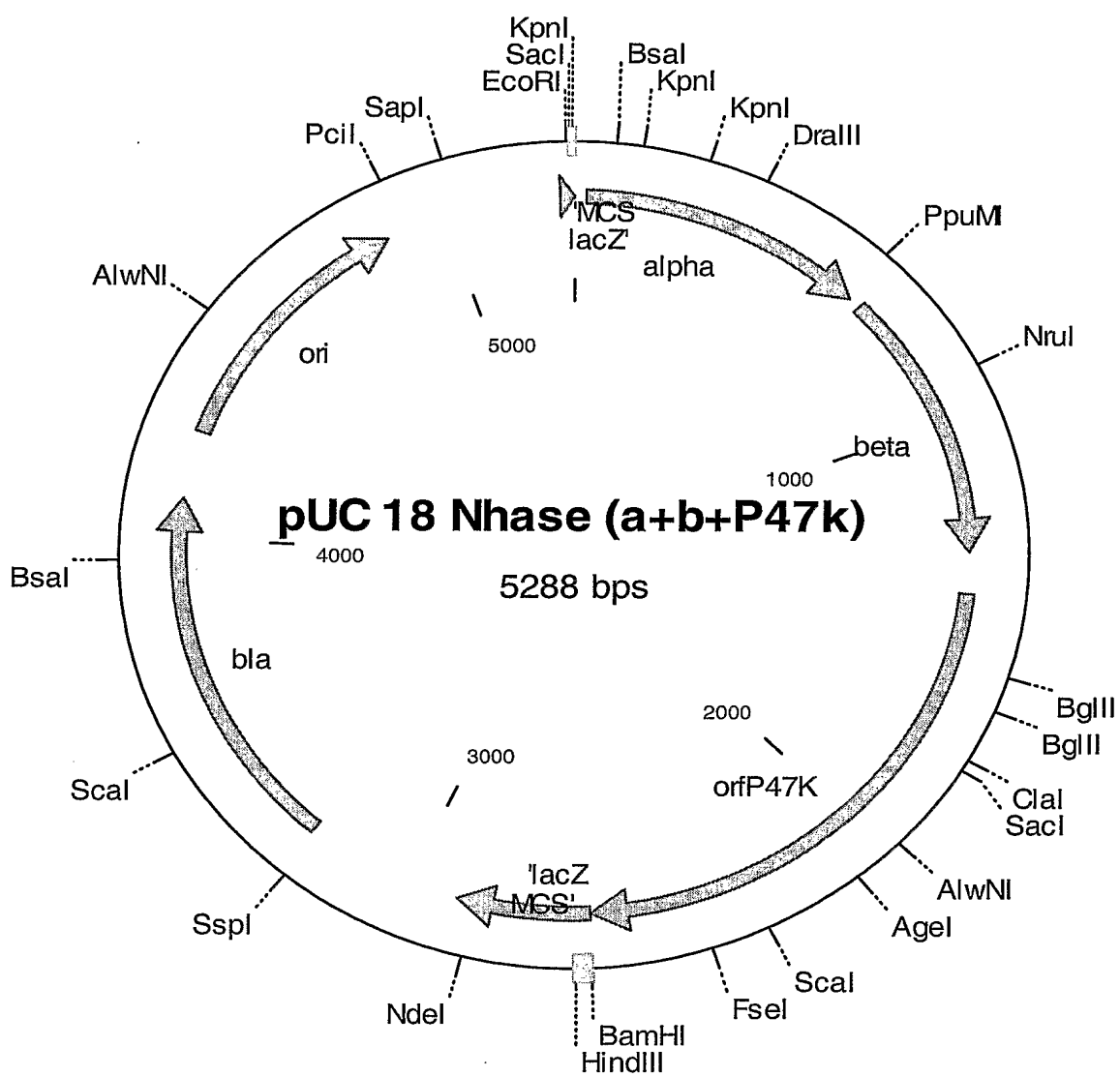


Abb. 2

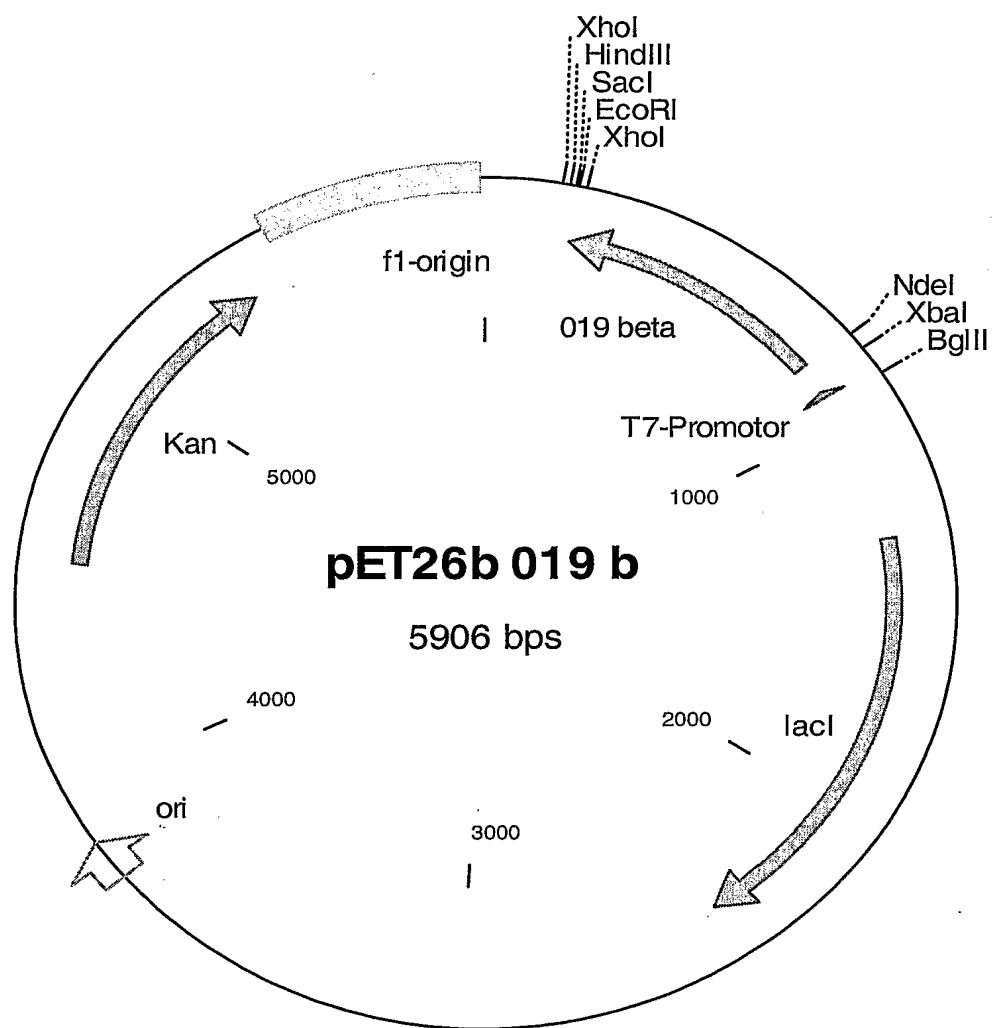


Abb. 3

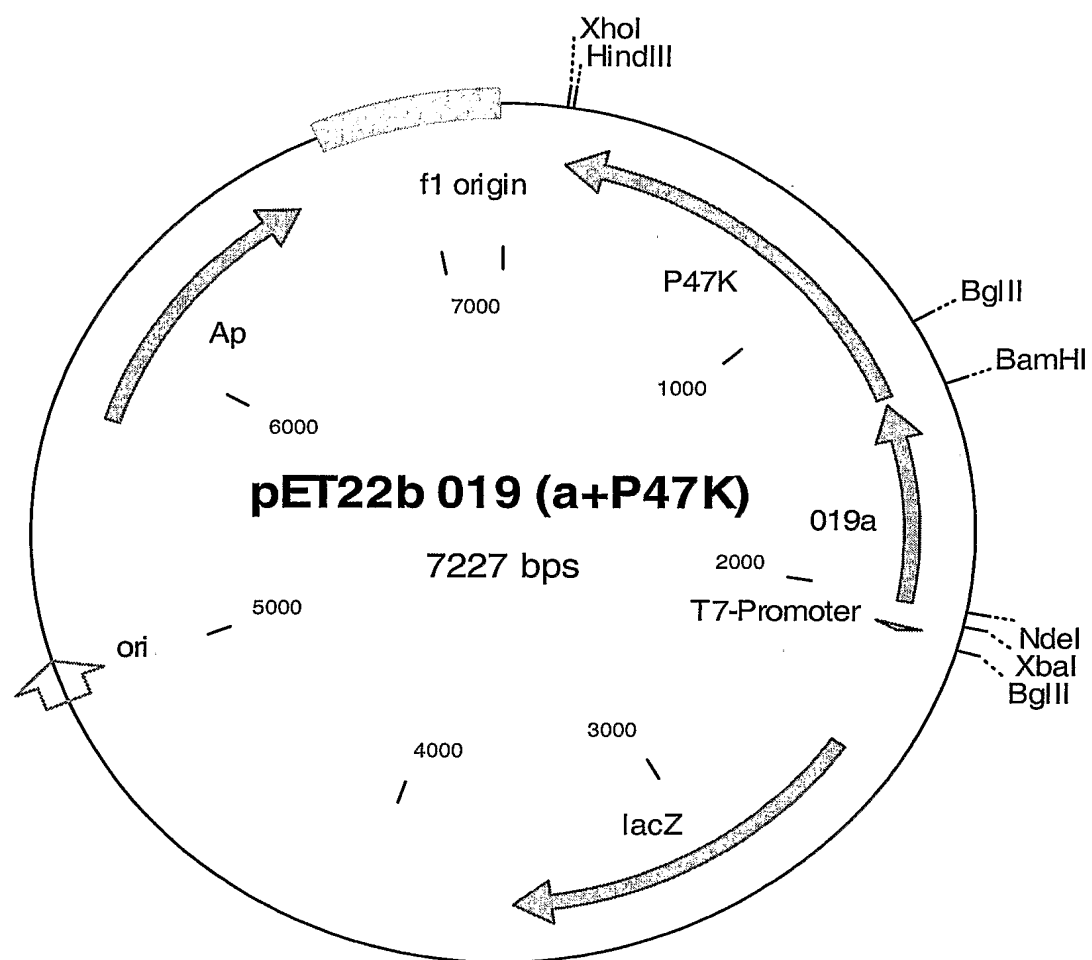


Abb. 4

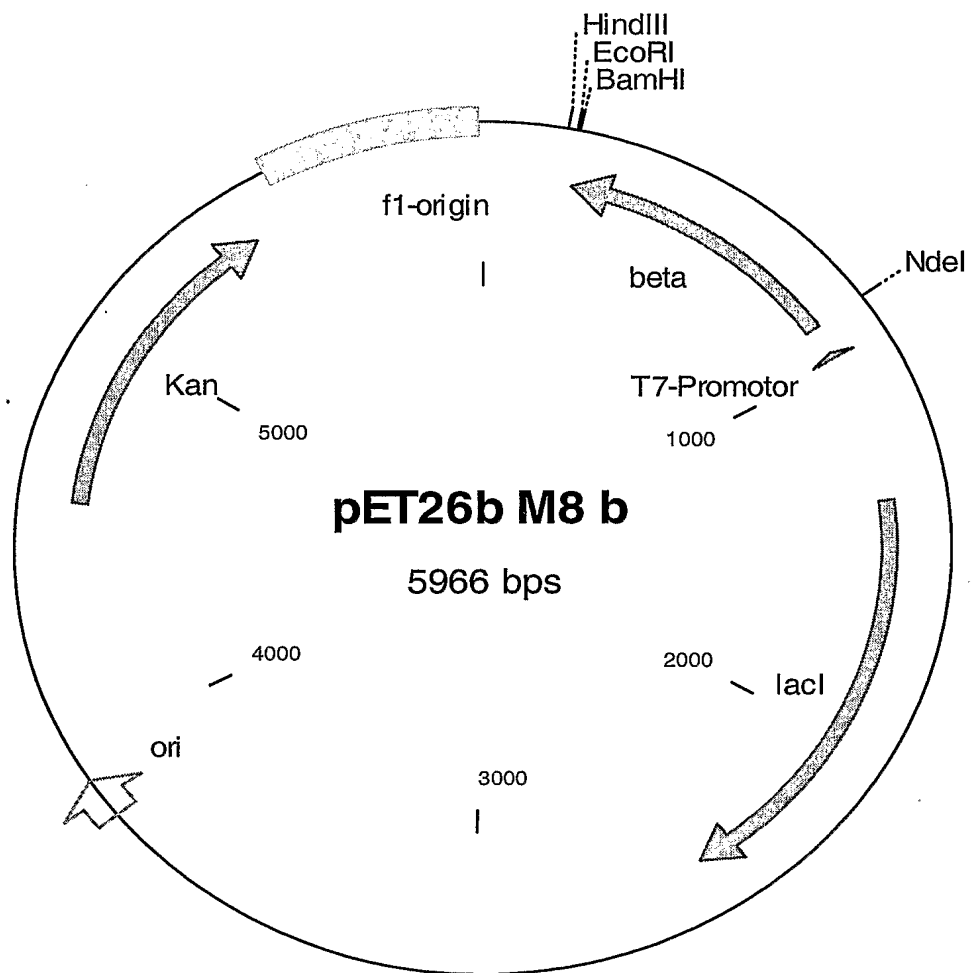
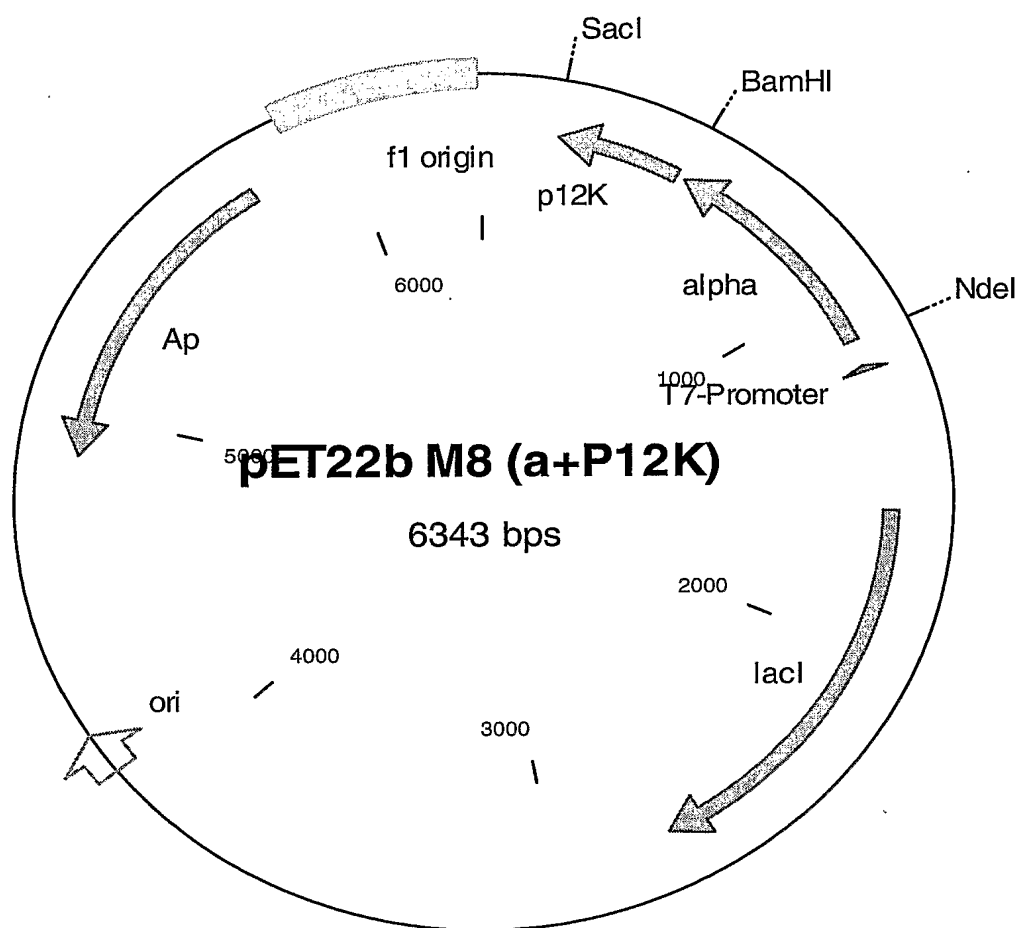


Abb. 5



SEQUENCE LISTING

<110> Degussa AG
 5 <120> Expression von Nitrilhydratasen im Zwei-Vektor-Expressionssystem
 <130> 040065 AM
 <160> 34
 10 <170> PatentIn version 3.1
 <210> 1
 <211> 624
 15 <212> DNA
 <213> Rhodococcus erythropolis
 <220>
 <221> CDS
 20 <222> (1) .. (624)
 <223>
 <400> 1
 25 atg tca gta acg atc gac cac aca acg gag aac gcc gca ccg gcc cag 48
 Met Ser Val Thr Ile Asp His Thr Thr Glu Asn Ala Ala Pro Ala Gln
 1 5 10 15
 gcg ccg gtc tcc gat cgc gcg tgg gcc ctg ttc cgc gca ctc gac ggt 96
 Ala Pro Val Ser Asp Arg Ala Trp Ala Leu Phe Arg Ala Leu Asp Gly
 30 20 25 30
 aag gga ttg gta ccc gac ggt tac gtc gag gga tgg aag aag acc ttc 144
 Lys Gly Leu Val Pro Asp Gly Tyr Val Glu Gly Trp Lys Lys Thr Phe
 35 35 40 45
 gag gag gac ttc agt cca agg cgc gga gcg gaa ttg gtc gcg cgg gcg 192
 Glu Glu Asp Phe Ser Pro Arg Arg Gly Ala Glu Leu Val Ala Arg Ala
 50 55 60
 40 tgg acc gac ccc gat ttc cgg caa ctg ctt ctc acc gac ggt acc gcc 240
 Trp Thr Asp Pro Asp Phe Arg Gln Leu Leu Leu Thr Asp Gly Thr Ala
 65 70 75 80
 gcg gtt gcc cag tac gga tat ctg ggc ccc cag ggc gaa tac atc gtg 288
 Ala Val Ala Gln Tyr Gly Tyr Leu Gly Pro Gln Gly Glu Tyr Ile Val
 45 85 90 95
 gca gtc gaa gac acc ccg acc ctc aag aac gtg atc gtg tgc tcg ctg 336
 Ala Val Glu Asp Thr Pro Thr Leu Lys Asn Val Ile Val Cys Ser Leu
 50 100 105 110
 tgt tca tgc acc gcg tgg ccc att ctc ggc ctg ccc cct acc tgg tac 384
 Cys Ser Cys Thr Ala Trp Pro Ile Leu Gly Leu Pro Pro Thr Trp Tyr
 115 120 125
 55 aag agt ttc gaa tac cgt gcg cga gtg gtg cgt gag cca cgg aag gtt 432
 Lys Ser Phe Glu Tyr Arg Ala Arg Val Val Arg Glu Pro Arg Lys Val
 130 135 140

ctc ttc gag atg gga acc gag atc gcg tcg gac gtc gag atc cgc gtc 480
 Leu Phe Glu Met Gly Thr Glu Ile Ala Ser Asp Val Glu Ile Arg Val
 145 150 155 160

5 tac gac acc acc gcc gaa act cgc tac atg gtt ctc ccg caa cgt ccc 528
 Tyr Asp Thr Thr Ala Glu Thr Arg Tyr Met Val Leu Pro Gln Arg Pro
 165 170 175

10 gca ggc acc gaa ggc tgg agc cag gaa cag ctt cag gag atc gtc acc 576
 Ala Gly Thr Glu Gly Trp Ser Gln Glu Gln Leu Gln Glu Ile Val Thr
 180 185 190

15 aag gac tgc ctg atc ggc gtc gca gtc ccg cag gtc ccc acc gtc tga 624
 Lys Asp Cys Leu Ile Gly Val Ala Val Pro Gln Val Pro Thr Val
 195 200 205

20 <210> 2
 <211> 207
 <212> PRT
 <213> Rhodococcus erythropolis
 <400> 2

25 Met Ser Val Thr Ile Asp His Thr Thr Glu Asn Ala Ala Pro Ala Gln
 1 5 10 15

30 Ala Pro Val Ser Asp Arg Ala Trp Ala Leu Phe Arg Ala Leu Asp Gly
 20 25 30

35 Lys Gly Leu Val Pro Asp Gly Tyr Val Glu Gly Trp Lys Lys Thr Phe
 35 40 45

40 Glu Glu Asp Phe Ser Pro Arg Arg Gly Ala Glu Leu Val Ala Arg Ala
 50 55 60

40 Trp Thr Asp Pro Asp Phe Arg Gln Leu Leu Leu Thr Asp Gly Thr Ala
 65 70 75 80

45 Ala Val Ala Gln Tyr Gly Tyr Leu Gly Pro Gln Gly Glu Tyr Ile Val
 85 90 95

50 Ala Val Glu Asp Thr Pro Thr Leu Lys Asn Val Ile Val Cys Ser Leu
 100 105 110

55 Cys Ser Cys Thr Ala Trp Pro Ile Leu Gly Leu Pro Pro Thr Trp Tyr
 115 120 125

Lys Ser Phe Glu Tyr Arg Ala Arg Val Val Arg Glu Pro Arg Lys Val
 130 135 140

Leu Phe Glu Met Gly Thr Glu Ile Ala Ser Asp Val Glu Ile Arg Val
 145 150 155 160

5 Tyr Asp Thr Thr Ala Glu Thr Arg Tyr Met Val Leu Pro Gln Arg Pro
 165 170 175

10 Ala Gly Thr Glu Gly Trp Ser Gln Glu Gln Leu Gln Glu Ile Val Thr
 180 185 190

15 Lys Asp Cys Leu Ile Gly Val Ala Val Pro Gln Val Pro Thr Val
 195 200 205

20 <210> 3
 <211> 639
 <212> DNA
 <213> Rhodococcus erythropolis

25 <220>
 <221> CDS
 <222> (1)..(639)
 <223>

30 <400> 3
 atg gat gga gta cac gat ctt gcc gga gtt caa ggc ttc ggc aaa gtc 48
 Met Asp Gly Val His Asp Leu Ala Gly Val Gln Gly Phe Gly Lys Val
 1 5 10 15

35 ccg cat acc gtc aac gcc gac atc ggc ccc acc ttc cac gcc gag tgg 96
 Pro His Thr Val Asn Ala Asp Ile Gly Pro Thr Phe His Ala Glu Trp
 20 25 30

40 gaa cac ctg ccg tac agc ctg atg ttc gcc ggt gtc gcc gaa ctc ggg 144
 Glu His Leu Pro Tyr Ser Leu Met Phe Ala Gly Val Ala Glu Leu Gly
 35 40 45

45 gca ttc agc gtc gac gaa gtt cga tac gtc gtc gag cgg atg gaa cca 192
 Ala Phe Ser Val Asp Glu Val Arg Tyr Val Val Glu Arg Met Glu Pro
 50 55 60

50 cgc cac tac atg atg acc ccg tac tac gag agg tac gtc atc ggc gtc 240
 Arg His Tyr Met Met Thr Pro Tyr Tyr Glu Arg Tyr Val Ile Gly Val
 65 70 75 80

55 gcg aca ctg atg gtc gaa aag gga atc ctg acg cag gaa gaa ctc gaa 288
 Ala Thr Leu Met Val Glu Lys Gly Ile Leu Thr Gln Glu Glu Leu Glu
 85 90 95

336 agc ctt gca ggg gga ccg ttc cca ctg tcg cgg ccc agc gaa tcc gaa
 Ser Leu Ala Gly Pro Phe Pro Leu Ser Arg Pro Ser Glu Ser Glu
 100 105 110

384 ggg cgg ccg gca ccc gtc gag acg acc acc ttc gaa atc ggt cag cga
 Gly Arg Pro Ala Pro Val Glu Thr Thr Thr Phe Glu Ile Gly Gln Arg
 115 120 125

gta cgc gtg cgc gac gag tac gtt ccg ggg cat att cga atg cct gcg 432
 Val Arg Val Arg Asp Glu Tyr Val Pro Gly His Ile Arg Met Pro Ala
 130 135 140

5

tac tgc cgc gga cga gtg gga acc atc tct cat cgg act acc gag aag 480
 Tyr Cys Arg Gly Arg Val Gly Thr Ile Ser His Arg Thr Thr Glu Lys
 145 150 155 160

10

tgg ccg ttt ccc gac gca atc ggc cac ggg cgc aac gac gcc ggc gaa 528
 Trp Pro Phe Pro Asp Ala Ile Gly His Gly Arg Asn Asp Ala Gly Glu
 165 170 175

15

gaa ccg acg tac cac gtg aag ttc gac gcc gag gaa ttg ttc ggt agc 576
 Glu Pro Thr Tyr His Val Lys Phe Asp Ala Glu Glu Leu Phe Gly Ser
 180 185 190

20

gac acc gac ggc ggc agc gtc gta gtc gac ctt ttc gag ggt tac ctc 624
 Asp Thr Asp Gly Gly Ser Val Val Val Asp Leu Phe Glu Gly Tyr Leu
 195 200 205

25

gag cct gcg gcc tga 639
 Glu Pro Ala Ala
 210

30

<210> 4
 <211> 212
 <212> PRT
 <213> Rhodococcus erythropolis
 <400> 4

35

Met Asp Gly Val His Asp Leu Ala Gly Val Gln Gly Phe Gly Lys Val
 1 5 10 15

40

Pro His Thr Val Asn Ala Asp Ile Gly Pro Thr Phe His Ala Glu Trp
 20 25 30

45

Glu His Leu Pro Tyr Ser Leu Met Phe Ala Gly Val Ala Glu Leu Gly
 35 40 45

50

Ala Phe Ser Val Asp Glu Val Arg Tyr Val Val Glu Arg Met Glu Pro
 50 55 60

55

Arg His Tyr Met Met Thr Pro Tyr Tyr Glu Arg Tyr Val Ile Gly Val
 65 70 75 80

Ala Thr Leu Met Val Glu Lys Gly Ile Leu Thr Gln Glu Glu Leu Glu
 85 90 95

Ser Leu Ala Gly Gly Pro Phe Pro Leu Ser Arg Pro Ser Glu Ser Glu
 100 105 110

5 Gly Arg Pro Ala Pro Val Glu Thr Thr Thr Phe Glu Ile Gly Gln Arg
 115 120 125
 Val Arg Val Arg Asp Glu Tyr Val Pro Gly His Ile Arg Met Pro Ala
 130 135 140
 10 Tyr Cys Arg Gly Arg Val Gly Thr Ile Ser His Arg Thr Thr Glu Lys
 145 150 155 160
 15 Trp Pro Phe Pro Asp Ala Ile Gly His Gly Arg Asn Asp Ala Gly Glu
 165 170 175
 20 Glu Pro Thr Tyr His Val Lys Phe Asp Ala Glu Glu Leu Phe Gly Ser
 180 185 190
 25 Asp Thr Asp Gly Gly Ser Val Val Val Asp Leu Phe Glu Gly Tyr Leu
 195 200 205
 Glu Pro Ala Ala
 210
 30 <210> 5
 <211> 624
 <212> DNA
 <213> Rhodococcus erythropolis
 35 <220>
 <221> CDS
 <222> (1)..(624)
 <223>
 40 <400> 5
 atg tca gta acg atc gac cac aca acg gag aac gcc gca ccg gcc cag 48
 Met Ser Val Thr Ile Asp His Thr Thr Glu Asn Ala Ala Pro Ala Gln
 1 5 10 15
 45 gcg ccg gtc tcc gat cgc gcg tgg gcc ctg ttc cgc gca ctc gac ggt 96
 Ala Pro Val Ser Asp Arg Ala Trp Ala Leu Phe Arg Ala Leu Asp Gly
 20 25 30
 50 aag gga ttg gta ccc gac ggt tac gtc gaa gga tgg aag aaa acc ttc 144
 Lys Gly Leu Val Pro Asp Gly Tyr Val Glu Gly Trp Lys Lys Thr Phe
 35 40 45
 55 gag gag gac ttc agt cca agg cgc gga gcg gaa ttg gtc gcg cgg gcg 192
 Glu Glu Asp Phe Ser Pro Arg Arg Gly Ala Glu Leu Val Ala Arg Ala
 50 55 60

	tgg acc gac ccc gag ttc cgg cag ttg ctt ctc acc gac ggt acc gcc	240
	Trp Thr Asp Pro Glu Phe Arg Gln Leu Leu Leu Thr Asp Gly Thr Ala	
	65 70 75 80	
5	gcg gtt gcc cag tac gga tac ctg ggc ccc cag ggc gag tac atc gtg	288
	Ala Val Ala Gln Tyr Gly Tyr Leu Gly Pro Gln Gly Glu Tyr Ile Val	
	85 90 95	
10	gca gtc gaa gac acc ccg acc ctc aag aac gtg atc gtg tgc tcg ctg	336
	Ala Val Glu Asp Thr Pro Thr Leu Lys Asn Val Ile Val Cys Ser Leu	
	100 105 110	
15	tgt tca tgc acc gcg tgg ccc att ctc ggc ctg ccc cct acc tgg tac	384
	Cys Ser Cys Thr Ala Trp Pro Ile Leu Gly Leu Pro Pro Thr Trp Tyr	
	115 120 125	
20	aag agt ttc gaa tac cgt gcg cga gtg gtg cgt gag cca cgg aag gtt	432
	Lys Ser Phe Glu Tyr Arg Ala Arg Val Val Arg Glu Pro Arg Lys Val	
	130 135 140	
	ctc tcc gag atg gga acc gag atc gcg tcg gac gtc gag atc cgc gtc	480
	Leu Ser Glu Met Gly Thr Glu Ile Ala Ser Asp Val Glu Ile Arg Val	
	145 150 155 160	
25	tac gac acc acc gcc gaa act cgc tac atg gtt ctc ccg caa cgt ccc	528
	Tyr Asp Thr Thr Ala Glu Thr Arg Tyr Met Val Leu Pro Gln Arg Pro	
	165 170 175	
30	gca ggc acc gaa ggc tgg agc cag gaa caa ctg cag gaa atc gtc acc	576
	Ala Gly Thr Glu Gly Trp Ser Gln Glu Gln Leu Gln Glu Ile Val Thr	
	180 185 190	
35	aag gac tgc ctg atc ggc gtc gca gtc ccg cag gtc ccc acc gtc tga	624
	Lys Asp Cys Leu Ile Gly Val Ala Val Pro Gln Val Pro Thr Val	
	195 200 205	
40	<210> 6	
	<211> 207	
	<212> PRT	
	<213> Rhodococcus erythropolis	
	<400> 6	
45	Met Ser Val Thr Ile Asp His Thr Thr Glu Asn Ala Ala Pro Ala Gln	
	1 5 10 15	
50	Ala Pro Val Ser Asp Arg Ala Trp Ala Leu Phe Arg Ala Leu Asp Gly	
	20 25 30	
55	Lys Gly Leu Val Pro Asp Gly Tyr Val Glu Gly Trp Lys Lys Thr Phe	
	35 40 45	
	Glu Glu Asp Phe Ser Pro Arg Arg Gly Ala Glu Leu Val Ala Arg Ala	
	50 55 60	

Trp Thr Asp Pro Glu Phe Arg Gln Leu Leu Leu Thr Asp Gly Thr Ala
 65 70 75 80

5

Ala Val Ala Gln Tyr Gly Tyr Leu Gly Pro Gln Gly Glu Tyr Ile Val
 85 90 95

10

Ala Val Glu Asp Thr Pro Thr Leu Lys Asn Val Ile Val Cys Ser Leu
 100 105 110

15

Cys Ser Cys Thr Ala Trp Pro Ile Leu Gly Leu Pro Pro Thr Trp Tyr
 115 120 125

20

Lys Ser Phe Glu Tyr Arg Ala Arg Val Val Arg Glu Pro Arg Lys Val
 130 135 140

Leu Ser Glu Met Gly Thr Glu Ile Ala Ser Asp Val Glu Ile Arg Val
 145 150 155 160

25

Tyr Asp Thr Thr Ala Glu Thr Arg Tyr Met Val Leu Pro Gln Arg Pro
 165 170 175

30

Ala Gly Thr Glu Gly Trp Ser Gln Glu Gln Leu Gln Glu Ile Val Thr
 180 185 190

35

Lys Asp Cys Leu Ile Gly Val Ala Val Pro Gln Val Pro Thr Val
 195 200 205

40

<210> 7
 <211> 639
 <212> DNA
 <213> Rhodococcus erythropolis

45

<220>
 <221> CDS
 <222> (1)..(639)
 <223>

50

<400> 7
 atg gat gga gta cac gat ctt gcc gga gtt caa ggc ttc ggc aaa gtc 48
 Met Asp Gly Val His Asp Leu Ala Gly Val Gln Gly Phe Gly Lys Val
 1 5 10 15

55

ccg cat acc gtc aac gcc gac atc ggc ccc acc ttc cac gcc gag tgg 96
 Pro His Thr Val Asn Ala Asp Ile Gly Pro Thr Phe His Ala Glu Trp
 20 25 30

gaa cac ctg ccg tac agc ctg atg ttc gcc ggt gtc gcc gaa ctc ggg 144
 Glu His Leu Pro Tyr Ser Leu Met Phe Ala Gly Val Ala Glu Leu Gly
 35 40 45

5 gca ttc agc gtc gac gaa gtt cga tac gtc gtc gag cgg atg gaa cca 192
 Ala Phe Ser Val Asp Glu Val Arg Tyr Val Val Glu Arg Met Glu Pro
 50 55 60
 10 cgc cac tac atg atg acc ccg tac tac gag agg tac gtc atc ggc gtc 240
 Arg His Tyr Met Met Thr Pro Tyr Tyr Glu Arg Tyr Val Ile Gly Val
 65 70 75 80
 15 gcg aca ctg atg gtc gaa aag gga atc ctg acg cag gat gaa ctc gaa 288
 Ala Thr Leu Met Val Glu Lys Gly Ile Leu Thr Gln Asp Glu Leu Glu
 85 90 95
 20 agc ctt gca ggg gga ccg ttc cca ctg tcg cgg ccc agc gaa tcc gaa 336
 Ser Leu Ala Gly Gly Pro Phe Pro Leu Ser Arg Pro Ser Glu Ser Glu
 100 105 110
 25 ggg cgt ccg gca ccc gtc gag acg acc acc ttc gaa atc ggt cag cga 384
 Gly Arg Pro Ala Pro Val Glu Thr Thr Thr Phe Glu Ile Gly Gln Arg
 115 120 125
 30 gta cgc gtg cgc gac gag tac gtt ccg ggg cat att cga atg cct gcg 432
 Val Arg Val Arg Asp Glu Tyr Val Pro Gly His Ile Arg Met Pro Ala
 130 135 140
 35 tac tgc cgc gga cga gtg gga acc atc tct cat cgg act acc gag aag 480
 Tyr Cys Arg Gly Arg Val Gly Thr Ile Ser His Arg Thr Thr Glu Lys
 145 150 155 160
 40 tgg cca ttt ccc gac gca atc ggc cac ggg cgc aac gac gcc ggc gaa 528
 Trp Pro Phe Pro Asp Ala Ile Gly His Gly Arg Asn Asp Ala Gly Glu
 165 170 175
 45 gaa ccg acg tac cac gtg aag ttc gcc gcc gag gaa ttg ttc ggt agc 576
 Glu Pro Thr Tyr His Val Lys Phe Ala Ala Glu Glu Leu Phe Gly Ser
 180 185 190
 50 gac acc gac ggc ggc agc gtc gta gtc gac ctt ttc gag ggt tac ctc 624
 Asp Thr Asp Gly Gly Ser Val Val Val Asp Leu Phe Glu Gly Tyr Leu
 195 200 205
 55 gag cct gcg gcc tga 639
 Glu Pro Ala Ala
 210
 <210> 8
 <211> 212
 <212> PRT
 <213> Rhodococcus erythropolis
 <400> 8
 Met Asp Gly Val His Asp Leu Ala Gly Val Gln Gly Phe Gly Lys Val
 1 5 10 15
 Pro His Thr Val Asn Ala Asp Ile Gly Pro Thr Phe His Ala Glu Trp
 20 25 30

5 Glu His Leu Pro Tyr Ser Leu Met Phe Ala Gly Val Ala Glu Leu Gly
 35 40 45

10 Ala Phe Ser Val Asp Glu Val Arg Tyr Val Val Glu Arg Met Glu Pro
 50 55 60

15 Arg His Tyr Met Met Thr Pro Tyr Tyr Glu Arg Tyr Val Ile Gly Val
 65 70 75 80

20 Ala Thr Leu Met Val Glu Lys Gly Ile Leu Thr Gln Asp Glu Leu Glu
 85 90 95

25 Ser Leu Ala Gly Gly Pro Phe Pro Leu Ser Arg Pro Ser Glu Ser Glu
 100 105 110

30 Gly Arg Pro Ala Pro Val Glu Thr Thr Thr Phe Glu Ile Gly Gln Arg
 115 120 125

35 Val Arg Val Arg Asp Glu Tyr Val Pro Gly His Ile Arg Met Pro Ala
 130 135 140

40 Tyr Cys Arg Gly Arg Val Gly Thr Ile Ser His Arg Thr Thr Glu Lys
 145 150 155 160

45 Trp Pro Phe Pro Asp Ala Ile Gly His Gly Arg Asn Asp Ala Gly Glu
 165 170 175

50 Glu Pro Thr Tyr His Val Lys Phe Ala Ala Glu Glu Leu Phe Gly Ser
 180 185 190

55 Asp Thr Asp Gly Gly Ser Val Val Val Asp Leu Phe Glu Gly Tyr Leu
 195 200 205

60 Glu Pro Ala Ala
 210

65 <210> 9
 <211> 624
 <212> DNA
 <213> Rhodococcus erythropolis

70 <220>
 <221> CDS
 <222> (1) .. (624)
 <223>

<400> 9
 atg tca gta acg atc gac cac aca acg gag aac gcc gca ccg gcc cag 48
 Met Ser Val Thr Ile Asp His Thr Thr Glu Asn Ala Ala Pro Ala Gln
 5 1 5 10 15

 gcg ccg gtc tcc gac cgg gcg tgg gcc ctg ttc cgc gca ctc gac ggt 96
 Ala Pro Val Ser Asp Arg Ala Trp Ala Leu Phe Arg Ala Leu Asp Gly
 20 25 30
 10
 aag gga ttg gta ccc gac ggt tac gtc gag gga tgg aag aag acc ttc 144
 Lys Gly Leu Val Pro Asp Gly Tyr Val Glu Gly Trp Lys Lys Thr Phe
 35 40 45
 15
 gag gag gac ttc agt cca agg cgc gga gcg gaa ttg gtc gcg cgg gcg 192
 Glu Glu Asp Phe Ser Pro Arg Arg Gly Ala Glu Leu Val Ala Arg Ala
 50 55 60
 20
 tgg acc gac ccc gag ttc cgg cag ttg ctt ctc acc gac ggt acc gcc 240
 Trp Thr Asp Pro Glu Phe Arg Gln Leu Leu Leu Thr Asp Gly Thr Ala
 65 70 75 80
 25
 gcg gtt gcc cag tac gga tat ctg ggc ccc cag ggc gag tac atc gtg 288
 Ala Val Ala Gln Tyr Gly Tyr Leu Gly Pro Gln Gly Glu Tyr Ile Val
 85 90 95
 30
 gca gtc gaa gac acc ccg acc ctc aag aac gtg atc gtg tgc tcg ttg 336
 Ala Val Glu Asp Thr Pro Thr Leu Lys Asn Val Ile Val Cys Ser Leu
 100 105 110
 35
 tgt tca tgc acc gcg tgg ccc att ctc ggc ctg ccc cct acc tgg tac 384
 Cys Ser Cys Thr Ala Trp Pro Ile Leu Gly Leu Pro Pro Thr Trp Tyr
 115 120 125
 40
 aag agt ttc gaa tac cgt gcg cga gtg gtg cgt gag cca ccg aag gtt 432
 Lys Ser Phe Glu Tyr Arg Ala Arg Val Val Arg Glu Pro Arg Lys Val
 130 135 140
 45
 ctc tcc gag atg gga acc gag atc gcg tcg gac gtc gag atc cgc gtc 480
 Leu Ser Glu Met Gly Thr Glu Ile Ala Ser Asp Val Glu Ile Arg Val
 145 150 155 160
 50
 tac gac acc acc gcc gaa act cgc tac atg gtt ctc ccg caa cgt ccc 528
 Tyr Asp Thr Thr Ala Glu Thr Arg Tyr Met Val Leu Pro Gln Arg Pro
 165 170 175
 55
 gca ggc acc gaa ggc tgg agc cag gaa cag ctt caa gag atc gtc acc 576
 Ala Gly Thr Glu Gly Trp Ser Gln Glu Gln Leu Gln Glu Ile Val Thr
 180 185 190
 60
 aag gac tgc ctg atc ggc gtc gca gtc ccg cag gtc ccc acc gtc tga 624
 Lys Asp Cys Leu Ile Gly Val Ala Val Pro Gln Val Pro Thr Val
 195 200 205
 65
 <210> 10
 <211> 207
 <212> PRT
 <213> Rhodococcus erythropolis

<400> 10

5 Met Ser Val Thr Ile Asp His Thr Thr Glu Asn Ala Ala Pro Ala Gln
 1 5 10 15
 Ala Pro Val Ser Asp Arg Ala Trp Ala Leu Phe Arg Ala Leu Asp Gly
 20 25 30
 10 Lys Gly Leu Val Pro Asp Gly Tyr Val Glu Gly Trp Lys Lys Thr Phe
 35 40 45
 15 Glu Glu Asp Phe Ser Pro Arg Arg Gly Ala Glu Leu Val Ala Arg Ala
 50 55 60
 20 Trp Thr Asp Pro Glu Phe Arg Gln Leu Leu Leu Thr Asp Gly Thr Ala
 65 70 75 80
 25 Ala Val Ala Gln Tyr Gly Tyr Leu Gly Pro Gln Gly Glu Tyr Ile Val
 85 90 95
 Ala Val Glu Asp Thr Pro Thr Leu Lys Asn Val Ile Val Cys Ser Leu
 100 105 110
 30 Cys Ser Cys Thr Ala Trp Pro Ile Leu Gly Leu Pro Pro Thr Trp Tyr
 115 120 125
 35 Lys Ser Phe Glu Tyr Arg Ala Arg Val Val Arg Glu Pro Arg Lys Val
 130 135 140
 40 Leu Ser Glu Met Gly Thr Glu Ile Ala Ser Asp Val Glu Ile Arg Val
 145 150 155 160
 45 Tyr Asp Thr Thr Ala Glu Thr Arg Tyr Met Val Leu Pro Gln Arg Pro
 165 170 175
 Ala Gly Thr Glu Gly Trp Ser Gln Glu Gln Leu Gln Glu Ile Val Thr
 180 185 190
 50 Lys Asp Cys Leu Ile Gly Val Ala Val Pro Gln Val Pro Thr Val
 195 200 205
 55 <210> 11
 <211> 639
 <212> DNA
 <213> Rhodococcus erythropolis

<220>
 <221> CDS
 <222> (1)..(639)
 5 <223>

<400> 11
 atg gat gga gta cac gat ctt gcc gga gtt caa ggc ttc ggc aaa gtc 48
 Met Asp Gly Val His Asp Leu Ala Gly Val Gln Gly Phe Gly Lys Val
 10 1 5 10 15

ccg cat acc gtc aac gcc gac atc ggc ccc acc ttc cac gcc gag tgg 96
 Pro His Thr Val Asn Ala Asp Ile Gly Pro Thr Phe His Ala Glu Trp
 20 20 25 30

gaa cac ctg ccg tac agc ctg atg ttc gcc ggt gtc gcc gaa ctc ggg 144
 Glu His Leu Pro Tyr Ser Leu Met Phe Ala Gly Val Ala Glu Leu Gly
 15 35 40 45

gca ttc agc gtc gac gaa gtt cga tac gtc gtc gag cgg atg gaa cca 192
 Ala Phe Ser Val Asp Glu Val Arg Tyr Val Val Glu Arg Met Glu Pro
 50 55 60

cgc cac tac atg atg acc ccg tac tac gag agg tac gtc atc ggc gtc 240
 Arg His Tyr Met Met Thr Pro Tyr Tyr Glu Arg Tyr Val Ile Gly Val
 25 65 70 75 80

gcg aca ctg atg gtc gaa aag gga atc ctg acg cag gaa gaa ctc gaa 288
 Ala Thr Leu Met Val Glu Lys Gly Ile Leu Thr Gln Glu Glu Leu Glu
 30 85 90 95

agc ctt gca ggg gga ccg ttc cca ctg tcg cgg cca agc gaa tcc gaa 336
 Ser Leu Ala Gly Gly Pro Phe Pro Leu Ser Arg Pro Ser Glu Ser Glu
 35 100 105 110

ggg cgt ccg gca ccc gtc gag acg acc acc ttc gaa gtc ggt cag cga 384
 Gly Arg Pro Ala Pro Val Glu Thr Thr Thr Phe Glu Val Gly Gln Arg
 115 120 125

gta cgc gtg cgc gac gag tac gtt ccg ggg cat att cga atg cct gcg 432
 Val Arg Val Arg Asp Glu Tyr Val Pro Gly His Ile Arg Met Pro Ala
 130 135 140

tac tgc cgc gga cga gtg gga acc atc tct cat cgg act acc gag aag 480
 Tyr Cys Arg Gly Arg Val Gly Thr Ile Ser His Arg Thr Thr Glu Lys
 145 150 155 160

tgg cca ttt ccc gac gca atc ggc cac ggg cgc aac gac gcc ggc gaa 528
 Trp Pro Phe Pro Asp Ala Ile Gly His Gly Arg Asn Asp Ala Gly Glu
 50 165 170 175

gaa ccg acg tac cac gtg aag ttc gac gcc gag gaa ttg ttc ggt agc 576
 Glu Pro Thr Tyr His Val Lys Phe Asp Ala Glu Glu Leu Phe Gly Ser
 180 185 190

gac acc gac ggc ggc agc gtc gta gtc gac ctt ttc gag ggt tac ctc 624
 Asp Thr Asp Gly Gly Ser Val Val Val Asp Leu Phe Glu Gly Tyr Leu
 195 200 205

gag cct gcg gcc tga
 Glu Pro Ala Ala
 210

5

<210> 12
 <211> 212
 <212> PRT
 <213> Rhodococcus erythropolis

10

<400> 12

15

Met Asp Gly Val His Asp Leu Ala Gly Val Gln Gly Phe Gly Lys Val
 1 5 10 15

Pro His Thr Val Asn Ala Asp Ile Gly Pro Thr Phe His Ala Glu Trp
 20 25 30

20

Glu His Leu Pro Tyr Ser Leu Met Phe Ala Gly Val Ala Glu Leu Gly
 35 40 45

25

Ala Phe Ser Val Asp Glu Val Arg Tyr Val Val Glu Arg Met Glu Pro
 50 55 60

30

Arg His Tyr Met Met Thr Pro Tyr Tyr Glu Arg Tyr Val Ile Gly Val
 65 70 75 80

35

Ala Thr Leu Met Val Glu Lys Gly Ile Leu Thr Gln Glu Glu Leu Glu
 85 90 95

40

Ser Leu Ala Gly Gly Pro Phe Pro Leu Ser Arg Pro Ser Glu Ser Glu
 100 105 110

Gly Arg Pro Ala Pro Val Glu Thr Thr Thr Phe Glu Val Gly Gln Arg
 115 120 125

45

Val Arg Val Arg Asp Glu Tyr Val Pro Gly His Ile Arg Met Pro Ala
 130 135 140

50

Tyr Cys Arg Gly Arg Val Gly Thr Ile Ser His Arg Thr Thr Glu Lys
 145 150 155 160

55

Trp Pro Phe Pro Asp Ala Ile Gly His Gly Arg Asn Asp Ala Gly Glu
 165 170 175

Glu Pro Thr Tyr His Val Lys Phe Asp Ala Glu Glu Leu Phe Gly Ser
 180 185 190

	Asp Thr Asp Gly Gly Ser Val Val Val Asp Leu Phe Glu Gly Tyr Leu	
	195 200 205	
5	Glu Pro Ala Ala	
	210	
10	<210> 13	
	<211> 612	
	<212> DNA	
	<213> Rhodococcus erythropolis	
15	<220>	
	<221> CDS	
	<222> (1)..(612)	
	<223>	
20	<400> 13	
	gtg agc gag cac gtc aat aag tac acg gag tac gag gca cgt acc aag	48
	Val Ser Glu His Val Asn Lys Tyr Thr Glu Tyr Glu Ala Arg Thr Lys	
	1 5 10 15	
25	gca atc gaa act ttg ctg tac gag cga ggg ctc atc acg ccc gcc gcg	96
	Ala Ile Glu Thr Leu Leu Tyr Glu Arg Gly Leu Ile Thr Pro Ala Ala	
	20 25 30	
30	gtc gac cga gtc gtt tcg tac tac gag aac gag atc ggc ccg atg ggc	144
	Val Asp Arg Val Val Ser Tyr Tyr Glu Asn Glu Ile Gly Pro Met Gly	
	35 40 45	
35	ggc gcc aag gtc gtg gcg aag tcc tgg gtg gac cct gag tac cgc aag	192
	Gly Ala Lys Val Val Ala Lys Ser Trp Val Asp Pro Glu Tyr Arg Lys	
	50 55 60	
40	tgg ctc gaa gag gac gcg acg gcc gcg atg gcg tca ttg ggc tat gcc	240
	Trp Leu Glu Glu Asp Ala Thr Ala Ala Met Ala Ser Leu Gly Tyr Ala	
	65 70 75 80	
45	ggc gag cag gca cac caa att tcg gcg gtc ttc aac gac tcc caa acg	288
	Gly Glu Gln Ala His Gln Ile Ser Ala Val Phe Asn Asp Ser Gln Thr	
	85 90 95	
50	cat cac gtg gtg gtg tgc act ctg tgt tcg tgc tat ccg tgg ccg gtg	336
	His His Val Val Val Cys Thr Leu Cys Ser Cys Tyr Pro Trp Pro Val	
	100 105 110	
55	ctt ggt ctc ccg ccc gcc tgg tac aag agc atg gag tac cgg tcc cga	384
	Leu Gly Leu Pro Pro Ala Trp Tyr Lys Ser Met Glu Tyr Arg Ser Arg	
	115 120 125	
60	gtg gta gcg gac cct cgt gga gtg ctc aag cgc gat ttc ggt ttc gac	432
	Val Val Ala Asp Pro Arg Gly Val Leu Lys Arg Asp Phe Gly Phe Asp	
	130 135 140	
65	atc ccc gat gag gtg gag gtc agg gtt tgg gac agc agc tcc gaa atc	480
	Ile Pro Asp Glu Val Glu Val Arg Val Trp Asp Ser Ser Ser Glu Ile	
	145 150 155 160	

5 cgc tac atc gtc atc ccg gaa cgg ccg gcc ggc acc gac ggt tgg tcc 528
 Arg Tyr Ile Val Ile Pro Glu Arg Pro Ala Gly Thr Asp Gly Trp Ser
 165 170 175

10 gag gac gag ctg gcg aag ctg gtg agt cgg gac tcg atg atc ggt gtc 576
 Glu Asp Glu Leu Ala Lys Leu Val Ser Arg Asp Ser Met Ile Gly Val
 180 185 190

15 agt aat gcg ctc aca ccc cag gaa gtg atc gta tga 612
 Ser Asn Ala Leu Thr Pro Gln Glu Val Ile Val
 195 200

20 <210> 14
 <211> 203
 <212> PRT
 <213> Rhodococcus erythropolis

25 <400> 14
 Val Ser Glu His Val Asn Lys Tyr Thr Glu Tyr Glu Ala Arg Thr Lys
 1 5 10 15

30 Ala Ile Glu Thr Leu Leu Tyr Glu Arg Gly Leu Ile Thr Pro Ala Ala
 20 25 30

35 Val Asp Arg Val Val Ser Tyr Tyr Glu Asn Glu Ile Gly Pro Met Gly
 35 40 45

40 Gly Ala Lys Val Val Ala Lys Ser Trp Val Asp Pro Glu Tyr Arg Lys
 50 55 60

45 Trp Leu Glu Glu Asp Ala Thr Ala Ala Met Ala Ser Leu Gly Tyr Ala
 65 70 75 80

50 Gly Glu Gln Ala His Gln Ile Ser Ala Val Phe Asn Asp Ser Gln Thr
 85 90 95

55 His His Val Val Val Cys Thr Leu Cys Ser Cys Tyr Pro Trp Pro Val
 100 105 110

60 Leu Gly Leu Pro Pro Ala Trp Tyr Lys Ser Met Glu Tyr Arg Ser Arg
 115 120 125

65 Val Val Ala Asp Pro Arg Gly Val Leu Lys Arg Asp Phe Gly Phe Asp
 130 135 140

70 Ile Pro Asp Glu Val Glu Val Arg Val Trp Asp Ser Ser Ser Glu Ile
 145 150 155 160

5 Arg Tyr Ile Val Ile Pro Glu Arg Pro Ala Gly Thr Asp Gly Trp Ser
 165 170 175
 10 Glu Asp Glu Leu Ala Lys Leu Val Ser Arg Asp Ser Met Ile Gly Val
 180 185 190
 15 Ser Asn Ala Leu Thr Pro Gln Glu Val Ile Val
 195 200
 20 <210> 15
 <211> 690
 <212> DNA
 <213> Rhodococcus erythropolis
 25 <220>
 <221> CDS
 <222> (1)..(690)
 <223>
 30 <400> 15
 atg gat ggt atc cac gac aca ggc ggc atg acc gga tac gga ccg gtc 48
 Met Asp Gly Ile His Asp Thr Gly Gly Met Thr Gly Tyr Gly Pro Val
 1 5 10 15
 35 ccc tat cag aag gac gag ccc ttc ttc cac tac gag tgg gag ggt cgg 96
 Pro Tyr Gln Lys Asp Glu Pro Phe Phe His Tyr Glu Trp Glu Gly Arg
 20 25 30
 40 acc ctg tcg att ctg acc tgg atg cat ctc aag ggc atg tcg tgg tgg 144
 Thr Leu Ser Ile Leu Thr Trp Met His Leu Lys Gly Met Ser Trp Trp
 35 40 45
 45 gac aag tcg cgg ttc ttc cgg gag tcg atg ggg aac gaa aac tac gtc 192
 Asp Lys Ser Arg Phe Phe Arg Glu Ser Met Gly Asn Glu Asn Tyr Val
 50 55 60
 50 aac gag att cgc aac tcg tac tac acc cac tgg ctg agt gcg gca gaa 240
 Asn Glu Ile Arg Asn Ser Tyr Tyr Thr His Trp Leu Ser Ala Ala Glu
 65 70 75 80
 55 cgt atc ctc gtc gcc gac aag atc atc acc gaa gaa gag cga aag cac 288
 Arg Ile Leu Val Ala Asp Lys Ile Ile Thr Glu Glu Glu Arg Lys His
 85 90 95
 60 cgt gtg cag gag atc ctc gag ggt cgg tac acg gac agg aac ccg tcg 336
 Arg Val Gln Glu Ile Leu Glu Gly Arg Tyr Thr Asp Arg Asn Pro Ser
 100 105 110
 65 cgg aag ttc gat ccg gcc gag atc gag aag gcg atc gaa cgg ctt cac 384
 Arg Lys Phe Asp Pro Ala Glu Ile Glu Lys Ala Ile Glu Arg Leu His
 115 120 125

gag ccc cac tcc cta gca ctt cca gga gcg gag ccg agt ttc tcc ctc 432
 Glu Pro His Ser Leu Ala Leu Pro Gly Ala Glu Pro Ser Phe Ser Leu
 130 135 140

5 ggt gac aag gtc aaa gtg aag aat atg aac ccg ctg gga cac aca cgg 480
 Gly Asp Lys Val Lys Val Lys Asn Met Asn Pro Leu Gly His Thr Arg
 145 150 155 160

10 tgc ccg aaa tat gtg cgg aac aag atc ggg gaa atc gtc acc tcc cac 528
 Cys Pro Lys Tyr Val Arg Asn Lys Ile Gly Glu Ile Val Thr Ser His
 165 170 175

15 ggc tgc cag atc tat ccc gag agc agc tcc gcc ggc ctc ggc gac gat 576
 Gly Cys Gln Ile Tyr Pro Glu Ser Ser Ser Ala Gly Leu Gly Asp Asp
 180 185 190

20 ccc cgc ccg ctc tac acg gtc gcg ttt tcc gcc cag gaa ctg tgg ggc 624
 Pro Arg Pro Leu Tyr Thr Val Ala Phe Ser Ala Gln Glu Leu Trp Gly
 195 200 205

gac gac gga aac ggg aaa gac gta gtg tgc gtc gat ctc tgg gaa ccg 672
 Asp Asp Gly Asn Gly Lys Asp Val Val Cys Val Asp Leu Trp Glu Pro
 210 215 220

25 tac ctg atc tct gcg tga 690
 Tyr Leu Ile Ser Ala
 225

30 <210> 16
 <211> 229
 <212> PRT
 <213> Rhodococcus erythropolis

35 <400> 16

Met Asp Gly Ile His Asp Thr Gly Gly Met Thr Gly Tyr Gly Pro Val
 1 5 10 15

40 Pro Tyr Gln Lys Asp Glu Pro Phe Phe His Tyr Glu Trp Glu Gly Arg
 20 25 30

45 Thr Leu Ser Ile Leu Thr Trp Met His Leu Lys Gly Met Ser Trp Trp
 35 40 45

50 Asp Lys Ser Arg Phe Phe Arg Glu Ser Met Gly Asn Glu Asn Tyr Val
 50 55 60

55 Asn Glu Ile Arg Asn Ser Tyr Tyr Thr His Trp Leu Ser Ala Ala Glu
 65 70 75 80

Arg Ile Leu Val Ala Asp Lys Ile Ile Thr Glu Glu Glu Arg Lys His
 85 90 95

Arg Val Gln Glu Ile Leu Glu Gly Arg Tyr Thr Asp Arg Asn Pro Ser
 100 105 110

5

Arg Lys Phe Asp Pro Ala Glu Ile Glu Lys Ala Ile Glu Arg Leu His
 115 120 125

10

Glu Pro His Ser Leu Ala Leu Pro Gly Ala Glu Pro Ser Phe Ser Leu
 130 135 140

15

Gly Asp Lys Val Lys Val Lys Asn Met Asn Pro Leu Gly His Thr Arg
 145 150 155 160

20

Cys Pro Lys Tyr Val Arg Asn Lys Ile Gly Glu Ile Val Thr Ser His
 165 170 175

Gly Cys Gln Ile Tyr Pro Glu Ser Ser Ser Ala Gly Leu Gly Asp Asp
 180 185 190

25

Pro Arg Pro Leu Tyr Thr Val Ala Phe Ser Ala Gln Glu Leu Trp Gly
 195 200 205

30

Asp Asp Gly Asn Gly Lys Asp Val Val Cys Val Asp Leu Trp Glu Pro
 210 215 220

35

Tyr Leu Ile Ser Ala
 225

40

<210> 17
 <211> 22
 <212> DNA
 <213> Artificial

45

<220>
 <223> Primer
 <400> 17
 gcccgcataa gaaaaggtga ac

50

<210> 18
 <211> 21
 <212> DNA
 <213> Artificial

55

<220>
 <223> Primer
 <400> 18
 gcatgccttc aaatcagcct g

5 <210> 19
 <211> 24
 <212> DNA
 <213> Artificial

10 <220>
 <223> Primer
 <400> 19
 agggtgaacc atatgtcagt aacg 24

15 <210> 20
 <211> 22
 <212> DNA
 <213> Artificial

20 <220>
 <223> Primer
 <400> 20
 tgtcggatcc atcagacggt gg 22

25 <210> 21
 <211> 23
 <212> DNA
 <213> Artificial

30 <220>
 <223> Primer

35 <400> 21
 agcaccatat ggatggagta cac 23

40 <210> 22
 <211> 21
 <212> DNA
 <213> Artificial

45 <220>
 <223> Primer
 <400> 22
 gttgggaatt caggccgcag g 21

50 <210> 23
 <211> 27
 <212> DNA
 <213> Artificial

55 <220>
 <223> Primer
 <400> 23

cgcggatcca agaaggagat atacatg 27

5 <210> 24
<211> 22
<212> DNA
<213> Artificial

10 <220>
<223> Primer

<400> 24
ccgcaacggtt caaacggtct gg 22

15 <210> 25
<211> 27
<212> DNA
<213> Artificial

20 <220>
<223> Primer

25 <400> 25
aggaatacgc atatgagcga gcacgtc 27

30 <210> 26
<211> 30
<212> DNA
<213> Artificial

35 <220>
<223> Primer

<400> 26
gtgtgggatcc actcatacga tcacttctctg 30

40 <210> 27
<211> 31
<212> DNA
<213> Artificial

45 <220>
<223> Primer

<400> 27
aggaatgagc atatggatgg tatccacgac a 31

50

55 <210> 28
<211> 33
<212> DNA
<213> Artificial

<220>
<223> Primer

<400> 28
 atcgggatcc ttccacgcag agatcaggta cgg 33

5 <210> 29
 <211> 35
 <212> DNA
 <213> Artificial

10 <220>
 <223> Primer

<400> 29
 15 ctcaggatcc aaggagtgat cgtatgagtg aagac 35

<210> 30
 <211> 26
 <212> DNA
 20 <213> Artificial

<220>
 <223> Primer

25 <400> 30
 acaggagctc tcagtcgatg atggcc 26

<210> 31
 30 <211> 315
 <212> DNA
 <213> Rhodococcus erythropolis

<220>
 35 <221> CDS
 <222> (1)..(315)
 <223>

<400> 31
 40 atg agt gaa gac aca ctc act gat cgg ctc ccg gcg act ggg acc gcc 48
 Met Ser Glu Asp Thr Leu Thr Asp Arg Leu Pro Ala Thr Gly Thr Ala
 1 5 10 15

gca ccg ccc cgc gac aat ggc gag ctt gta ttc acc gag cct tgg gaa 96
 45 Ala Pro Pro Arg Asp Asn Gly Glu Leu Val Phe Thr Glu Pro Trp Glu
 20 25 30

gca acg gca ttc ggg gtc gcc atc gcg ctt tcg gat cag aag tcg tac 144
 50 Ala Thr Ala Phe Gly Val Ala Ile Ala Leu Ser Asp Gln Lys Ser Tyr
 35 40 45

gaa tgg gag ttc ttc cga cag cgt ctc att cac tcc atc gct gag gcc 192
 Glu Trp Glu Phe Phe Arg Gln Arg Leu Ile His Ser Ile Ala Glu Ala
 50 55 60

55 aac ggt tgc gag gca tac tac gag agc tgg aca aag gcg ctc gag gcc 240
 Asn Gly Cys Glu Ala Tyr Tyr Glu Ser Trp Thr Lys Ala Leu Glu Ala
 65 70 75 80

agc gtg gtc gac tcg ggg ctg atc agc gaa gat gag atc cgc gag cgc 288
 Ser Val Val Asp Ser Gly Leu Ile Ser Glu Asp Glu Ile Arg Glu Arg
 85 90 95

5 atg gaa tcg atg gcc atc atc gac tga 315
 Met Glu Ser Met Ala Ile Ile Asp
 100

10 <210> 32
 <211> 104
 <212> PRT
 <213> Rhodococcus erythropolis

15 <400> 32
 Met Ser Glu Asp Thr Leu Thr Asp Arg Leu Pro Ala Thr Gly Thr Ala
 1 5 10 15

20 Ala Pro Pro Arg Asp Asn Gly Glu Leu Val Phe Thr Glu Pro Trp Glu
 20 25 30

25 Ala Thr Ala Phe Gly Val Ala Ile Ala Leu Ser Asp Gln Lys Ser Tyr
 35 40 45

30 Glu Trp Glu Phe Phe Arg Gln Arg Leu Ile His Ser Ile Ala Glu Ala
 50 55 60

35 Asn Gly Cys Glu Ala Tyr Tyr Glu Ser Trp Thr Lys Ala Leu Glu Ala
 65 70 75 80

40 Ser Val Val Asp Ser Gly Leu Ile Ser Glu Asp Glu Ile Arg Glu Arg
 85 90 95

45 Met Glu Ser Met Ala Ile Ile Asp
 100

45 <210> 33
 <211> 1200
 <212> DNA
 <213> Rhodococcus erythropolis

50 <220>
 <221> CDS
 <222> (1)..(1200)
 <223>

55 <400> 33
 atg gtc gac aca cga ctt ccg gtc acg gtg ctg tca ggt ttc ctg ggc 48
 Met Val Asp Thr Arg Leu Pro Val Thr Val Leu Ser Gly Phe Leu Gly
 1 5 10 15

	gcc ggg aag acg aca cta ctc aac gag atc ctg cga aat cga gag ggt	96
	Ala Gly Lys Thr Thr Leu Leu Asn Glu Ile Leu Arg Asn Arg Glu Gly	
	20 25 30	
5	cgg cgg gtc gcg gtg atc gtc aac gac atg agc gaa atc aac atc gac	144
	Arg Arg Val Ala Val Ile Val Asn Asp Met Ser Glu Ile Asn Ile Asp	
	35 40 45	
10	agt gca gaa gtc gag cgt gag atc tcg ctc agt cgc tcc gag gag aaa	192
	Ser Ala Glu Val Glu Arg Glu Ile Ser Leu Ser Arg Ser Glu Glu Lys	
	50 55 60	
15	ctg gtc gag atg acc aac ggc tgc atc tgc tgc act ctg cga gag gat	240
	Leu Val Glu Met Thr Asn Gly Cys Ile Cys Cys Thr Leu Arg Glu Asp	
	65 70 75 80	
20	ctt ctt tcc gag atc agc gcc ttg gcc gcc gat ggc cga ttc gac tac	288
	Leu Leu Ser Glu Ile Ser Ala Leu Ala Ala Asp Gly Arg Phe Asp Tyr	
	85 90 95	
25	cta ctc atc gaa tct tcg ggc atc tcc gaa ccg ctt ccc gtc gca gag	336
	Leu Leu Ile Glu Ser Ser Gly Ile Ser Glu Pro Leu Pro Val Ala Glu	
	100 105 110	
30	acg ttc aca ttc atc gat acc gac ggc cac gcc ctc gcc gac gtc gcc	384
	Thr Phe Thr Phe Ile Asp Thr Asp Gly His Ala Leu Ala Asp Val Ala	
	115 120 125	
35	cga ctc gac acc atg gtc acc gtc gtc gac ggc cac agt ttt ctg cgc	432
	Arg Leu Asp Thr Met Val Thr Val Val Asp Gly His Ser Phe Leu Arg	
	130 135 140	
40	gac tac acg gct ggg ggc cgc gtc gaa gcc gat gcc ccg gaa gac gaa	480
	Asp Tyr Thr Ala Gly Gly Arg Val Glu Ala Asp Ala Pro Glu Asp Glu	
	145 150 155 160	
45	cga gac atc gcg gat ctg ctt gtc gat cag atc gaa ttt gcc gac gtc	528
	Arg Asp Ile Ala Asp Leu Leu Val Asp Gln Ile Glu Phe Ala Asp Val	
	165 170 175	
50	atc ctg gtg agc aag gcc gat ctc gtc tcg cac cag cac ctg gtc gaa	576
	Ile Leu Val Ser Lys Ala Asp Leu Val Ser His Gln His Leu Val Glu	
	180 185 190	
55	ttg acc gca gtc ctg cgc tct ttg aac gca tcc gct gcg ata gtt ccg	624
	Leu Thr Ala Val Leu Arg Ser Leu Asn Ala Ser Ala Ala Ile Val Pro	
	195 200 205	
60	atg acg ctc ggt cgc atc cca ctc gac acg att ctc gac acc ggt ttg	672
	Met Thr Leu Gly Arg Ile Pro Leu Asp Thr Ile Leu Asp Thr Gly Leu	
	210 215 220	
65	ttc tcg ctc gaa aag gct gca cag gcc ccc gga tgg tta caa gaa ctc	720
	Phe Ser Leu Glu Lys Ala Ala Gln Ala Pro Gly Trp Leu Gln Glu Leu	
	225 230 235 240	
70	caa ggt gaa cac atc ccc gaa acc gaa gag tac gga atc agt tcg gtg	768
	Gln Gly Glu His Ile Pro Glu Thr Glu Glu Tyr Gly Ile Ser Ser Val	
	245 250 255	

	gtg tac cgc gag cgc gca ccc ttc cac ccc caa cgg ctg cat gat ttc	816
	Val Tyr Arg Glu Arg Ala Pro Phe His Pro Gln Arg Leu His Asp Phe	
	260 265 270	
5	ctc agc agc gag tgg acc aac gga aag tta ctt cgg gcc aag ggc tac	864
	Leu Ser Ser Glu Trp Thr Asn Gly Lys Leu Leu Arg Ala Lys Gly Tyr	
	275 280 285	
10	tac tgg aat gcc ggc cgg ttc acc gag atc ggg agt att tct cag gcc	912
	Tyr Trp Asn Ala Gly Arg Phe Thr Glu Ile Gly Ser Ile Ser Gln Ala	
	290 295 300	
15	ggt cat ctc att cgc cac gga tac gtc ggc cgt tgg tgg aag ttt cta	960
	Gly His Leu Ile Arg His Gly Tyr Val Gly Arg Trp Trp Lys Phe Leu	
	305 310 315 320	
20	ccc cgt gac gag tgg ccg gcc gac gat tac cgt cgt gac gga atc ctc	1008
	Pro Arg Asp Glu Trp Pro Ala Asp Asp Tyr Arg Arg Asp Gly Ile Leu	
	325 330 335	
25	gac aag tgg gaa gaa ccc gtc gga gac tgc cga caa gaa ctc gtc ttc	1056
	Asp Lys Trp Glu Glu Pro Val Gly Asp Cys Arg Gln Glu Leu Val Phe	
	340 345 350	
30	atc ggc caa gcc atc gac ccg tct cga ctg cac cga gaa ctc gac gcg	1104
	Ile Gly Gln Ala Ile Asp Pro Ser Arg Leu His Arg Glu Leu Asp Ala	
	355 360 365	
35	tgt cta ctc acc aca gcc gag atc gaa ctc ggg cca gac gtg tgg acc	1152
	Cys Leu Leu Thr Thr Ala Glu Ile Glu Leu Gly Pro Asp Val Trp Thr	
	370 375 380	
40	acc tgg agc gac ccc ctg ggc gtc ggc tat acc gac cag acc gtt tga	1200
	Thr Trp Ser Asp Pro Leu Gly Val Gly Tyr Thr Asp Gln Thr Val	
	385 390 395	
45	<210> 34	
	<211> 399	
	<212> PRT	
	<213> Rhodococcus erythropolis	
50	<400> 34	
55	Met Val Asp Thr Arg Leu Pro Val Thr Val Leu Ser Gly Phe Leu Gly	
	1 5 10 15	
60	Ala Gly Lys Thr Thr Leu Leu Asn Glu Ile Leu Arg Asn Arg Glu Gly	
	20 25 30	
65	Arg Arg Val Ala Val Ile Val Asn Asp Met Ser Glu Ile Asn Ile Asp	
	35 40 45	
70	Ser Ala Glu Val Glu Arg Glu Ile Ser Leu Ser Arg Ser Glu Glu Lys	
	50 55 60	

5 Leu Val Glu Met Thr Asn Gly Cys Ile Cys Cys Thr Leu Arg Glu Asp
 65 70 75 80

10 Leu Leu Ser Glu Ile Ser Ala Leu Ala Ala Asp Gly Arg Phe Asp Tyr
 85 90 95

15 Leu Leu Ile Glu Ser Ser Gly Ile Ser Glu Pro Leu Pro Val Ala Glu
 100 105 110

20 Thr Phe Thr Phe Ile Asp Thr Asp Gly His Ala Leu Ala Asp Val Ala
 115 120 125

25 Arg Leu Asp Thr Met Val Thr Val Val Asp Gly His Ser Phe Leu Arg
 130 135 140

30 Asp Tyr Thr Ala Gly Gly Arg Val Glu Ala Asp Ala Pro Glu Asp Glu
 145 150 155 160

35 Arg Asp Ile Ala Asp Leu Leu Val Asp Gln Ile Glu Phe Ala Asp Val
 165 170 175

40 Ile Leu Val Ser Lys Ala Asp Leu Val Ser His Gln His Leu Val Glu
 180 185 190

45 Leu Thr Ala Val Leu Arg Ser Leu Asn Ala Ser Ala Ala Ile Val Pro
 195 200 205

50 Met Thr Leu Gly Arg Ile Pro Leu Asp Thr Ile Leu Asp Thr Gly Leu
 210 215 220

55 Phe Ser Leu Glu Lys Ala Ala Gln Ala Pro Gly Trp Leu Gln Glu Leu
 225 230 235 240

60 Gln Gly Glu His Ile Pro Glu Thr Glu Glu Tyr Gly Ile Ser Ser Val
 245 250 255

65 Val Tyr Arg Glu Arg Ala Pro Phe His Pro Gln Arg Leu His Asp Phe
 260 265 270

70 Leu Ser Ser Glu Trp Thr Asn Gly Lys Leu Leu Arg Ala Lys Gly Tyr
 275 280 285

Tyr Trp Asn Ala Gly Arg Phe Thr Glu Ile Gly Ser Ile Ser Gln Ala
290 295 300

5 Gly His Leu Ile Arg His Gly Tyr Val Gly Arg Trp Trp Lys Phe Leu
305 310 315 320

10 Pro Arg Asp Glu Trp Pro Ala Asp Asp Tyr Arg Arg Asp Gly Ile Leu
325 330 335

15 Asp Lys Trp Glu Glu Pro Val Gly Asp Cys Arg Gln Glu Leu Val Phe
340 345 350

Ile Gly Gln Ala Ile Asp Pro Ser Arg Leu His Arg Glu Leu Asp Ala
355 360 365

20 Cys Leu Leu Thr Thr Ala Glu Ile Glu Leu Gly Pro Asp Val Trp Thr
370 375 380

25 Thr Trp Ser Asp Pro Leu Gly Val Gly Tyr Thr Asp Gln Thr Val
385 390 395